

The difference between science and dogma

Scientists and science teachers can draw useful lessons from the Kansas Board of Education's efforts to expel Charles Darwin from the state's schools.

In removing the teaching of evolution, as well as aspects of cosmology, from its curriculum guidance for school students, the Kansas Board of Education has done no favours to the schoolchildren of Kansas or to the national and international reputation of their state. It is to be hoped that the county boards that actually administer education in Kansas will ignore the board's guidelines, and that science teachers there will continue to appraise students of the nature of our Universe to the best of their ability — even if some of this knowledge is no longer to be tested in exams. The board may, however, have done the scientific community a small favour in the long run, if the right lessons are drawn from its decision.

The decision is the latest episode in a protracted effort by creationists in the United States to restrict discussion of evolution in the classroom. After attempts to ban the teaching of evolution outright were blocked in the courts, and efforts to encourage the teaching of 'creation science' met with equally limited success, they now seek to gently remove evolution from school curricula. The Kansas board has chosen to modify a set of school science teaching guidelines, proposed to it by a panel of science educators, by deleting almost two pages that refer to the theory of evolution. References to the Big Bang are also excised from the document.

It would be easy to mock the creationist idea, as some Americans have wasted no time in doing, or to highlight the weakness of the school board's apparent view that phenomena occur only if they can be directly observed. It is also all too easy to bemoan the polling data suggesting that the US public supports the creationist point of view — although one study suggests that US scientists feel the same way (*Nature* 386, 435; 1997). A more constructive response for scientists and science teachers is to ask themselves what they should be doing to keep the sentiment expressed in Kansas from getting out of hand.

One obvious course of action is for scientists to take a more active role in public life. This advice has been repeated to the point of tedium by Neal Lane, former head of the National Science Foundation and now President Bill Clinton's science adviser, and others. It has been repeated again this week by the American Geophysical Union (see page 701).

Another lesson to be gleaned from this episode is the omnipresent danger inherent in the teaching or presentation of science as a set of facts, to be learned by rote, rather than as, as it were, evolving body of knowledge. The National Academy of Sciences, in developing its own guidelines for school science in the United States, has consistently demanded a far greater emphasis on the scientific method itself.

If more children were taught science as a means of interrogating nature, rather than as a toolbox of rules, the 'debate' between creationism and Darwinism would come to be seen in a clearer light. Darwin's theory will continue to be attacked and will develop on the strength of new information established by geneticists, palaeontologists and others. The idea of divine creation is unlikely to benefit from such further investigation because it is not a scientific theory, but a dogma.

Phillip Johnson, a professor of law at the University of California at Berkeley and the author of *Darwin on Trial*, told *The New York Times* last week that the defence of evolution was becoming "the science educators' Vietnam". But the enlightenment is not going to be extinguished anytime soon. A better parallel would be with the Kansas skirmishes that preceded the American Civil War, when pro-slave factions from Missouri and the South sought to rout anti-slave Yankees from their settlement at Lawrence, now home to the University of Kansas. The initial rout was thoroughly successful, but, for a variety of reasons, the outcome of the wider conflict was never seriously in doubt. □

No time to hide

Britain is right to reject shrouding field trials of genetically modified crops in secrecy.

The UK government took a brave step earlier this week when it decided to make known publicly the precise locations of four new trial sites for tests of genetically modified (GM) crops, despite the increasing numbers of attacks on such sites by environmental activists. The four farm-scale trials of GM oilseed rape will begin this autumn at sites in Lincolnshire, Nottinghamshire and Hertfordshire, and will bring the total number of test sites to 75 next year.

The announcement, from the Department of the Environment, Transport and the Regions, quickly came under fire from farmers and commercial groups, who argued that it was an open invitation to protesters to destroy the trials (significantly, groups such as Greenpeace have declined to promise not to take such action). The British Plant Breeders Association, for example, says that it had recommended to the government that only the counties in which the trials were being held should be revealed, as is the practice in Germany and France.

But given that the issues at stake in the conflict over GM crops are as much to do with trust in public institutions as they are with rational action — or even scientific data — the government's decision in favour of openness must be applauded. The nuclear industry has already suffered the consequences of combining rigid secrecy with overstated claims to safety; the one can all too easily feed on the other, and attempts to break this particular vicious circle are therefore welcome.

They are also timely. In a week that has seen the company Monsanto heavily criticized by the UK Advertising Standards Authority for exaggerating some claims about the safety of its products in a public relations blitz in Britain last summer, the need to restore public credibility in the whole handling of the GM issue remains high on the political agenda (see page 702). Openness will not achieve this on its own. But the alternative is doomed to fail. □



Annotation competition spurs *Drosophila* sequencing efforts

[HEIDELBERG] Bioinformaticists have used a competition between rival research teams to identify the computational modelling tools that can most accurately locate and describe the genes hidden within a stretch of sequenced genome.

The results of the competition were assessed last week at the seventh international conference on Intelligent Systems for Molecular Biology (ISMB) in Heidelberg, Germany. The idea had been to stimulate the bioinformatics community to improve the predictive power of its software tools.

The immediate goal was to help identify the best tools for annotating the *Drosophila* genome, the sequence of which is due to be completed by the end of the year.

The idea of a competition came from Martin Reese, a postdoctoral fellow with the Berkeley *Drosophila* Genome Project (BDGP). He modelled it on the Critical Assessment of Techniques for Protein Structure Prediction (CASP) project, a biennial competition which asks computational modellers to predict the three-dimensional structure of proteins from DNA sequences.

Every other year, CASP competition organizers select proteins that are sequenced and whose structures are confidently expected to have been determined by X-ray crystallography by the time the competition ends (see *Nature Structural Biology* 6, 108; 1999).

In recognition of CASP, Reese has named his own project GASP — Genome Annotation Assessment Project. He says he wanted to “see how well our science community could work together”. But, unlike CASP, he says, the annotation project cannot be considered a real competition as there is no absolute ‘right’ answer, so no winners will be announced.

Earlier this year, the BDGP reached an agreement with the genomics company Celera to collaborate on sequencing and annotating the *Drosophila* genome, to which Celera is applying its ‘shotgun’ approach (see *Nature* 393, 296; 1998). Celera is a joint venture between the Perkin Elmer Corporation and The Institute for Genomic Research, directed by Craig Venter.

The BDGP is taking a slower but more thorough approach to the sequencing, and has already confirmed 25 megabases of the approximately 140-megabase genome. The Berkeley team will help Celera piece together its jumble of data. Celera got off to a late start because the delivery of 3700 DNA Analyzers produced by Applied Biosystems, part of Perkin Elmer, was delayed. Celera says it will begin making raw data public in October.



Warming up: Celera executives Mark Adams (right) and Hamilton Smith in the company's sequencing lab. High-quality software is vital for accurately sifting genes out of raw sequence data.

In May, Reese and BDGP project leader Gerald Rubin chose three megabases of their own unpublished sequence for the competition. Their own annotation, identifying and locating 220 genes in the stretch, was used as a reference. The competition was judged by an international panel of geneticists and bioinformaticists.

Rubin says he was “very pleasantly surprised” that 12 research teams around the world took up the challenge. Participants, who were given only six weeks to respond, say that they were equally happy to have the opportunity to benchmark their software tools against those of other groups.

One participant, Terry Gaasterland, a bioinformaticist at Rockefeller University in New York, says that the competition was “an amazing motivation for our group,” which had not used its software tools to approach such a complex organism. “We wanted to win, even knowing there is no real winner,” she says. “We would certainly do it again.”

One unexpected result, says Reese, is a consensus that the Berkeley group may have underestimated the number of genes in the stretch. Most participants predicted 20 or 30 more genes than the group had found.

Over the next six months, biologists at Berkeley will check the authenticity of all the predicted genes by seeing if the relevant sequences really do direct the synthesis of proteins in laboratory experiments.

“Annotation helps drive biological experiments,” says Rubin. And bioinformaticists find the feedback from experiments — “the reality check” — very encouraging, he says.

Equally importantly, the information about the quality of the software products being developed by different groups around

the world will aid efforts to annotate the *Drosophila* genome by highlighting the best tools. “The competition has been an excellent warm-up for the larger task of annotating the whole genome, and setting standards for the completeness and quality of the annotation,” says Mark Adams, Celera's vice-president for genome programmes.

For Celera and many BDGP scientists, the *Drosophila* genome is only a prelude to the human genome, whose full sequence Celera, engaged in sharp competition with publicly funded sequencing efforts, expects to complete by the end of 2001. “Our work on *Drosophila* will provide feedback for the Human Genome Project on how well shotgun sequencing works and on how best to annotate,” says Rubin.

The annotation competition will probably be repeated at a future ISMB conference. Reflecting the growing importance of informatics in biology, the ISMB is trying to cope with an almost geometric increase in attendance at meetings in recent years.

Thomas Lengauer, an institute director at the German Research Centre for Information Technology in Bonn, is delighted with the growth of interest. But he hopes that weight of numbers will not force organizers to hold more informal meetings where presentations are not refereed in advance. At Heidelberg, 35 papers were selected for presentation from the 150 submitted.

The ISMB will be sponsored in future by a new umbrella organization, the International Society for Computational Biology. Russ Altman, a bioinformaticist from Stanford University in California, was last week chosen as the society's president-elect, and Lengauer its vice-president-elect.

Alison Abbott

EPA science 'overburdened by Congress'

[WASHINGTON] Scientific research at the Environmental Protection Agency (EPA) consists of "a discontinuous series of broad, shallow efforts" according to an assessment commissioned by the EPA's research office. This is blamed on a combination of over-reaching congressional mandates and short-lived efforts at reform by the agency's political leadership.

The report was prepared by Resources for the Future, an independent think-tank, in an attempt to explain the often-cited shortcomings in EPA's scientific programme. Almost since its foundation in 1973, critics have charged that the regulatory agency does not do the high-quality research needed to frame its regulations.

Mark Powell, the study's author, says that

"the primary impediment" to research at the EPA "is that it isn't a science agency, it is first and foremost a regulatory agency". As the report notes: "The EPA's primary constituencies tend—with some justification—to view science and analysis as an obstacle to regulatory action."

Although the dominance of the agency by lawyers and other non-scientists is damaging to its science programmes, Powell argues that the largest culprit is the stream of regulatory mandates from the Congress which the agency has to implement. "New responsibilities have been added, but you never see any of the old ones being removed," he says.

Many of the mandates require the agency to re-assess acceptable levels of pollutants at regular intervals of three or five years, result-

ing in short bursts of research, which, the report says, "often result in a cyclic pattern of shallow, episodic initiatives".

Science has also suffered, however, because of its failure to deliver the results which the agency needs and which proponents of research have occasionally promised. Some analysts have suggested that scientists "in their eagerness for a place at the policy table, inadvertently over-sold risk assessment over a number of years as the elixir for ailing environmental policies," the report says.

It adds: "unfortunately, the basic scientific tools of environmental regulation, toxicology, and epidemiology are simply inadequate to provide precise answers regarding the effects of pollutants at low ambient levels".

Powell notes that when the EPA was established, almost one-third of its budget was devoted to its Office of Research and Development (ORD), but that this quickly fell back to less than 10 per cent and now stands at 5.8 per cent. When its programme offices are included, the agency is thought to spend about 8 per cent of its budget on research and development.

Resources for the Future's report says that this should be doubled to 15 per cent, and that special provision should be made for the ORD and the programme offices to plan and jointly support 'mid-term' research. It says that although the ORD supports basic research in universities, and programme offices work to meet immediate EPA mandates, work that can be applied in three or five years time—such as the development of new analysis tools and methods—is often neglected.

The report says that the agency's scientific programmes tend to be championed by the political leadership of the agency, but shunned by its senior permanent staff. "It is the political-level leadership which has been primarily responsible for driving the change toward increased use of science," it quotes one senior official as saying. But the report adds that "in general, scientists have little stature and power within EPA".

An EPA official said that it had not had much time to review the report, which was due to be delivered this week. The official added, however, that its conclusions may not reflect recent efforts by the ORD to expand scientific peer review and develop a new external grant programme, Science to Achieve Results.

But Powell is pessimistic that the report will instigate change. Some officials at the EPA agree with its recommendations, he says, "but there is some despondency that any sort of strategic planning will be derailed, either by the Congress or by political appointees".

Colin Macilwain

Clinton gets the message on biomass energy

[WASHINGTON] President Bill Clinton last week announced plans to boost research on biomass energy sources, with a view to tripling the use of biomass fuels by 2010. These currently account for 3 per cent of US energy use.

Speaking at the US Department of Agriculture on 12 August, Clinton said that biomass would provide extra farm income while enabling the United States to cut greenhouse-gas emissions. "Bio-energy is a means to heat our homes, to fuel our vehicles, to power our factories, while producing virtually no greenhouse-gas pollution," he said.

Clinton announced a scientific task force headed by Bill Richardson, the energy secretary, and Dan Glickman, the agriculture secretary, to draw up a research plan by the end of this year to prepare to meet the target. The task force is expected to recommend significant expansion of the \$200 million that the United States spends each year on biomass research.

Biomass advocates say that advances in biotechnology—in particular the availability of cheaply-produced enzymes that break down the cellulose in plants—are set to transform



Mass observation: studying trees for 'coppicing'.

the economics of biomass as an energy source.

The President's Council of Advisors on Science and Technology recently recommended increases in research into new energy sources, including biomass. But Clinton said he was alerted to the potential of biomass energy by a "brilliant" article in *Foreign Affairs* by Richard Lugar (Republican, Indiana), chair of the Senate Agriculture Committee, and James Woolsey, former head of the Central Intelligence Agency. The president added: "It's not only brilliant, but a guy who is scientifically challenged like me can understand it."

Henry Kelly, assistant director for technology at the White House Office of Science and Technology, says: "The science says that the time is ripe for a major

initiative" in biomass research.

The existing US ethanol programme, which subsidizes the conversion of corn seed into ethanol fuel, has been roundly criticized as an agriculture subsidy programme. Critics point out that the intensive farming of the corn fed into the ethanol plants consumes more energy than the fuel that comes out of them.

That would change if cellulose in the rest of the corn plant, which is currently wasted, or in other crops such as switchgrass, could be cheaply broken down. According to Lee Lynd, an engineer at Dartmouth College, New Hampshire, research could find enzymes that will do this cheaply, perhaps reproducing themselves as they go.

Clinton's announcement comes during a summer in which low rainfall and depressed crop prices are creating a major farm crisis in the United States. Officials said that the announcement would do little to help farmers in the near term. But Clinton saw it as an opportunity to argue that action to counter global warming could provide economic benefits, in this case for farmers.

C.M.

Kansas kicks evolution out the classroom

[SAN DIEGO] US scientists are being urged to become more involved in debates over the teaching of evolution following last week's decision by the Kansas Board of Education to halt the study of evolutionary biology in public schools.

Siding with creationists, the publicly elected Kansas board voted to approve standards for kindergarten through twelfth-grade students that eliminate the required teaching of evolution, cosmology and some sciences related to the Earth's origin. Local school districts will use these standards to set their curriculums.

"Many scientists were dismayed when the Kansas board voted to eliminate the theory of evolution from the state's science curriculum, and some will undoubtedly express shock and disapproval," says Fred Spilhaus, executive director of the 35,000-member American Geophysical Union (AGU). "What is needed from scientists now is not expressions of outrage, but active participation in state and local decision making."

But, while the AGU issued a rallying call, some major US organizations — including the American Association for the Advancement of Science and the National Academy of Sciences — declined to comment on the decision. The organizations have long-standing policies supporting the teaching of evolutionary biology.

In Kansas, there was shock and embarrassment. Governor Bill Graves, a Republican from a rural city, issued a statement saying: "This is a tragic, terrible, embarrassing solution to a problem that didn't exist."

The presidents of the six public universities in Kansas have said in a public statement

that the changed standards "will be detrimental to the future of science education in Kansas" and "will set Kansas back a century," giving "hard-to-find science teachers no choice but to pursue other career fields".

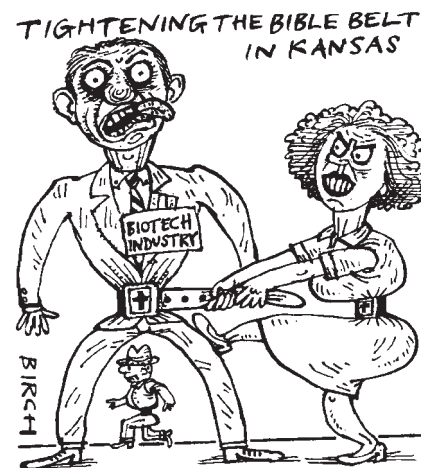
There are concerns that students will perform poorly in college entrance examinations, that universities will find it harder to attract quality science professors, and that businesses may have difficulty recruiting from other regions.

Steven Case, an experienced Kansas educator and head of Citizens for Science, a group that opposed the board's decision, said he feared the decision will have long-term negative consequences. Not teaching evolutionary biology, Case said, would impair the learning process, as studies show that knowledge is most likely to be retained "when the learner knows and understands underlying principles that can be applied to problems in new contexts".

The board "not only ignores good science, it also ignores what we know about learning," said Case. But Linda Holloway, president of the Kansas Board of Education, who voted in the six to four majority on 11 August, said previous state standards placed too "heavy an emphasis on evolution".

A former teacher of disabled children, Holloway from Shawnee, Kansas, says she "doesn't see evidence" for 'macroevolution' — the evolution of higher species from lower forms of life. But she says that she believes in 'microevolution' — the view that species can evolve within their group.

Critics of the board's new standards, says Holloway, "are making the assumption that local people don't have the capability to



decide what good education is on their own".

The new standards appear to be part of a campaign by creationists to further their cause through the political process. A Missouri-based group called the Creation Science Association of Mid-America, which opposes teaching evolution in schools, was closely involved in the Kansas effort.

Having put their imprint on state regulations in Kansas, the creationists are expected to move to the school-district level, and try to alter the school science curriculum.

Citizens for Science have said that the approved standards "are illegal to teach in public schools" and are the result of "a fringe group" following "a political strategy".

It remains unclear whether the battle will end up in court, as has occurred in other states. Kansas' governor wants a state constitutional amendment to abolish the board. But previous statewide votes to eliminate the board have failed.

Rex Dalton

Cray sale threatens to send researchers shopping elsewhere

[SAN DIEGO] The parent corporation of Cray supercomputers announced plans last week to sell the computer manufacturer, prompting new concern among scientific users already frustrated by slow part availability and the failure to develop competitive products.

The sale is part of a reorganization by Silicon Graphics Inc., the Mountain View, California, company that bought Cray about three years ago. Silicon Graphics also plans to lay off up to 1,500 employees, and concentrate on developing mid-sized, high-performance servers for the Internet.

The Cray supercomputers and some associated technologies will form a new business unit under the direction of Steve Oberlin, a respected supercomputer specialist. Silicon Graphics says that it then plans to sell all or part of that unit.

"We have decided that pursuing these

agendas separately but in coordination creates the best opportunity for the success of both businesses," says Rick Belluzzo, Silicon Graphics' chairman and chief executive.

But the news met with scepticism and concern among some Cray users, who have endured what they call years of marginal service. "The sun is setting on Cray," says one. "And it is sad." Some clients are already switching to other brands of computer.

The Cray vector supercomputers, made in Eagan, Minnesota, and Chippewa Falls, Wisconsin, were developed by the late Seymour Cray. Researchers value them for their ability to produce data rapidly from each cycle of the computer's internal clock. They are good for climate and weather forecasting, and have long been used by the US National Weather Service in Maryland.

But supercomputer specialists say that,

in recent years, the Cray has fallen behind its competitors. Many of Cray's best computer scientists have left the company, moving to competitors such as IBM, which now is taking Cray's clients.

The National Weather Service is replacing its Cray C90 with an IBM SPRS/6000. "The IBM beat out Cray on performance for the price," said Carl Staton, director of central operations for the Weather Service. And the San Diego Supercomputer Center at the University of California at San Diego has bought an IBM Teraflops machine to augment its Crays.

Privately, one computer scientist said that purchasers of high-end equipment became frustrated with Cray's constantly changing strategic direction under Silicon Graphics. "They didn't have good direction," says the scientist. "Their strategic plan was different every six months."

R. D.

Research to benefit from cheaper Landsat images

[WASHINGTON] The US government is about to keep a seven-year-old promise by offering new Landsat images of the Earth at low prices, following a failed attempt to commercialize remote-sensing data in the 1980s and 1990s.

The Landsat 7 satellite, launched in April, will also provide much more thorough coverage of the Earth than its predecessors, with increased spatial resolution.

Project managers at the US space agency NASA and the US Geological Survey (USGS), which initially will share responsibility for operating the satellite, have validated its onboard systems and had hoped to begin routine image collection by 15 August.

But technical problems with outside users connecting to the Earth Observing System (EOS) data gateway have pushed back the date, probably to late August or early September, according to Ralph Thompson, programme manager at USGS for Landsat 7.

Like previous versions, Landsat 7 will return multispectral images with a resolution of 30 metres. But it adds a panchromatic channel with a resolution of 15 metres and a thermal infrared channel with a resolution of 60 metres. Users will be able to combine the high-resolution black-and-white images with the lower-resolution data to produce synthetic colour images with greater detail.

The satellite marks the return of Landsat to government ownership. In 1992, Congress admitted that attempts to privatize Landsat had failed and passed a law calling for the government to operate the next satellite in the series and to sell its data at cost price.

But the advent of low-cost imagery was delayed by the loss of Landsat 6 shortly after its launch in 1993, and by problems with the power-supply hardware of the Enhanced Thematic Mapper on Landsat 7 before the launch of the latter.

Now that Landsat 7 is operating in orbit as planned, images will be much cheaper. The Earth Observation Satellite Company (EOSAT), which marketed Landsat images in the mid-1980s, charged up to \$4,400 for a single 185 by 170 kilometre 'scene'.

Even a modest land-use study could therefore rack up a hefty bill for pictures. "You really couldn't afford it," says Samuel Goward, chairman of the University of Maryland geography department, who heads NASA's 14-member Landsat science team.

Raw Landsat 7 images will be sold through the USGS EROS Data Center in South Dakota for just \$475 a scene. This should put Landsat imagery in the hands of many more researchers, says Goward.



Cape Canaveral from above: Landsat 7 will capture around 450 high-quality images every day.

Equally important, he says, is the programme's commitment to collecting cloud-free images of the entire land surface of the Earth, with regular seasonal updating. Targets of interest can be photographed every 16 days, the period between Landsat passes over a particular location.

Around 150,000 Landsat 7 scenes will be collected each year — 250 every day at the EROS Data Center, and another 200 at the dozen or so international Landsat stations. These will store images of regional interest and set their own pricing policies. International users will also be able to buy images from the EROS centre.

Commercial software packages and more powerful desktop computers have combined to make the data easier to interpret and manipulate, he says. But he estimates that "we're only halfway" to the point where an untrained scientist can download a Landsat image and start using it for research.

According to Goward, a decade of high prices and patchy data availability (EOSAT had reduced the number of pictures taken by Landsat to save costs) have slowed the development of easy-to-use image analysis packages.

He adds that the Landsat science team and others can develop software to bridge the gap between the relatively small group accustomed to using remote-sensing images and the wider scientific community.

Thomas Lillesand, director of the University of Wisconsin at Madison's Environmental Remote Sensing Center and current president of the American Society for Photogrammetry and Remote Sensing, says the drastic reduction of Landsat data costs is partly a recognition that the commercial market in remote sensing is about to shift to higher-resolution data.

Tony Reichhardt

Monsanto rapped for misleading press advertisements

[LONDON] A costly publicity campaign by the US biotechnology company Monsanto has backfired in Britain with the finding by the British Advertising Standards Authority (ASA) that the company had wrongly depicted facts about genetically modified (GM) foods in its press advertisements.

Activist organizations and individuals — including the Green Party, GeneWatch and the Soil Association — filed 13 complaints against the full-page advertisements, which ran last summer in a number of UK newspapers. The ASA upheld four of the complaints.

In two advertisements featuring GM potatoes, the ASA found that the text mistakenly suggested that the potatoes had been approved in 20 countries including the United Kingdom. In another instance, the company asserts that it has tested the safety of its GM foods for 20 years. The ASA said that evidence for testing dates back only 16 years.

The ASA also took issue with an advertisement claiming that GM tomatoes require less pesticide, as well as Monsanto's claim that cross-species gene transference is merely an extension of traditional cross-breeding.

Alex Woolfall, a Monsanto spokesperson, says the goal of the campaign was to educate the public about GM foods and present its side of the debate. "We were trying a form of advertising that we knew might be controversial, because the subject is controversial."

Also last week, the European Commission postponed a committee meeting intended to decide whether to approve three new GM crops — a genetically engineered fodder beet from Monsanto and two strains of rapeseed from AgrEvo.

But following the tightening of regulations by member states in late June, few GM crops are expected to be approved. It could take up to two years for the new regulations come into effect, because they require the blessing of the European Parliament (see *Nature* 400, 7; 1999).

Commenting last week, the new US ambassador to the European Union, Richard Morningstar, writes in his *Letter from Brussels* that the *de facto* moratorium is at the centre of tension between the United States and Europe concerning GM foods. Europe's failure to approve bioengineered products will cost the United States \$200 million in corn sales this year.

Foreshadowing what could amount to a trade war between the United States and Europe, Morningstar says: "Until the EU can credibly separate science-based risk assessment and regulations from the political process, the outlook for resolution of this issue is bleak."

Heather McCabe

Japanese university plan under fire again

[TOKYO] Shigehiko Hasumi, the outspoken president of Tokyo University, has strongly criticized a government plan to turn Japan's 'national' universities into semi-autonomous agencies with greater administrative independence.

Hasumi, who is also chairman of the Association of National Universities, said at a press conference last week that the government's proposal would be "unsuitable for the university's management, education and research". He added that a plan to introduce evaluation by the Management and Coordination Agency — which oversees government administration — "would further restrict" the universities' freedom.

Hasumi's statements came after a committee of researchers and academics set up by the Ministry of Education, Science, Sports and Culture (Monbusho) decided it was important to increase competition between universities, even at the risk of threatening research and teaching at some of them.

The 99 national universities have been targeted by government reforms aimed at improving administrative efficiency by

restructuring government ministries and agencies (see *Nature* **389**, 897; 1997).

Changes have already been proposed for institutes attached to science-related ministries and for the 14 basic research institutes under Monbusho, including the Institute of Space and Astronautical Sciences (ISAS) and the High Energy Accelerator Research Organization (see *Nature* **398**, 272; 1999).

The government had postponed its plan to restructure national universities because of fierce opposition from the academic community and Monbusho (see *Nature* **395**, 730; 1998). But such moves have gathered momentum recently as the result of strong pressure within the government to reduce the number of civil servants by giving agency status to government-run organizations.

In a bid to reduce tension over the move, Monbusho recently set up an ad hoc committee to discuss how the universities might be reformed without affecting the standard of research and education.

The committee — whose members include Minoru Oda, former director of ISAS, Hiroo Imura, former dean of Kyoto

University, and Leo Esaki, former dean of Tsukuba University — held its first meeting last week. It concluded that the current government proposals, which emphasize performance targets related to costs, could jeopardize research and teaching at universities.

But they agreed that some competition must be introduced, perhaps by targeting funding at research groups and institutions with good track records, as opposed to Monbusho's current system, which disperses funds to universities without adequate consideration of their performance.

The committee hopes to agree on a general direction by the autumn and reach a final decision by next summer. But sources close to Monbusho say the ministry sees the creation of agencies as 'inevitable', and that the final decision is likely to be imposed on universities by the government whether they like it or not.

"The administrative reform plan is being forced upon us by the bureaucrats," said Hasumi. "But a serious conflict [between the government and universities] would be inevitable at some point." **Asako Saegusa**

Australia plans research reforms but without extra funding

[SYDNEY] David Kemp, Australia's minister of education, training and youth affairs, has proposed some major changes in the way research is organized, including giving incentives to some universities to identify "specialized niches" for themselves in research and teaching.

His strategy is based on the universities obtaining extra revenue by commercializing their research more effectively and expanding their research contracts with the private sector.

The government claims its scheme would strengthen research links with industry, which it considers to be a top priority. But Peter Cullen, president of the Federation of Australian Scientific and Technological Societies, describes forcing universities to seek more funding from business as an "unlikely" strategy. "Governments have tried to persuade business to lift their R&D without any success. What new levers they think the universities have is a puzzle."

Under the proposals, the Australian Research Council (ARC) — earlier rumoured to be facing a major reduction in its powers — would see its influence grow. It would be given enhanced status, with its own act of parliament and a stronger role in giving 'strategic' advice to government and administering grants worth A\$436 million (US\$284 million).

ARC's programmes for funding research would be consolidated and become "more



Kemp: wants stronger links with industry.

flexible". Vicki Sara, a biomedical scientist who currently chairs ARC, would see her full-time post replaced by a part-time position and a full-time chief executive.

Sara welcomes the new "independence" of ARC and the "affirmation that excellence must be the hallmark of research training". But most universities are reserving judgement

until they have established what the changes will mean for them.

To the concern of those who argue that the main problem facing Australian universities is funding, Kemp stresses that his top priority is to get the strategy right, with resource issues "to be considered down the track".

Brian Anderson, president of the Australian Academy of Science, approves of the "streamlining" of ARC, but warns that it will be beneficial only "if it's accompanied by additional funding". Some of the large, older universities, which dominate ARC grants, fear a considerable loss of support.

The 'group of eight' universities, which win around four-fifths of competitive grants,

is opening an office in Canberra to promote and protect their position.

Ministers of the coalition government have ordered an unrelenting squeeze on academia over four budgets. The May 1999 budget, for example, cut 'targeted research' by 5.5 per cent and 'other higher education R&D' by 2 per cent.

John Niland, president of the Australian Vice-Chancellors' Committee, calculates that the group of eight, of which his own University of New South Wales is a member, faces an overall loss of A\$17 million. But some smaller, regional universities believe they will get more funding for research.

Niland says that the proposal hides a centralization of direction and control of research, and thinks universities will have to do more paperwork — generating binding strategic research plans, for example — to justify their block funding for research. He says this could either "restrict the diversity" of research or be "little more than a time-consuming paper shuffle".

Kemp says his personal priority is the funding of all postgraduate students through scholarships that they can use at the university of their choice or take to another.

Within existing funding, Kemp claims that hundreds of millions of dollars will be saved by limiting the period of masters to two years and doctorates to three-and-a-half years. Universities have four months to comment on his proposals. **Peter Pockley**

Rutherford Appleton tipped to house UK/French synchrotron

[LONDON] UK government officials are reported to have recommended that the country's new synchrotron facility, to be built jointly with France, should be located at the Rutherford Appleton Laboratory south of Oxford (see *Nature* **400**, 489; 1999).

According to the newsletter *Research Fortnight*, a decision on the location has been delayed by the science minister, Lord David Sainsbury. He wants to allow time to consider the impact on the Daresbury Laboratory outside Manchester, which operates the current national Synchrotron Radiation Source (SRS).

Within the synchrotron community, Rutherford is considered the most logical choice of site for the new X-ray source, known as Diamond. However some believe the RAL site is a compromise between Daresbury and a site outside Cambridge.

The project's main funder, the Wellcome Trust, is said to have been pushing for siting the synchrotron within its Hinxton Hall genome campus, and the resulting protracted negotiations to have caused delays to the project. But siting Diamond at Rutherford would not necessarily mean the

closure of Daresbury.

The government has refused to say when a decision on a site for Diamond will be made. But the research council which operates Rutherford and Daresbury is expecting a decision in early September.

Japan's universities face independent scrutiny

[TOKYO] Japan's Ministry of Education, Science, Sports and Culture is planning to create an independent body to evaluate teaching and research at national universities. It is hoped that this will help prospective students to select universities on merit instead of by reputation.

The ministry also plans to reflect the results of external evaluations in the allocation of research funds, in line with the government's recent shift towards competitive and peer-reviewed grant schemes. The new body, to be set up next year, will seek external reviewers from private and national universities and industry.

Australian Senate looks at global warming

[SYDNEY] The Australian Senate announced last week that it is to launch an inquiry into "the progress and adequacy of Australia's

policies to reduce global warming". The inquiry will be carried out by a Senate committee, and will look at policies to reduce greenhouse-gas emissions, and potential improvements to these policies.

Australia was one of the strongest critics of mandatory emissions limits in the run-up to the Kyoto climate conference in 1997 (see *Nature* **390**, 323; 1997). But concern has been growing, with reports on the potential impact on coral reefs. The inquiry will consider the projected impact of climate change on coral reef systems, alpine areas and wetlands.

Biomedicine convention into law in five countries

[LONDON] The first internationally binding legal text designed to protect individuals against the misuse of biological and medical advances will come into force in five European countries in December. The Convention on Human Rights and Biomedicine will become legally binding in Denmark, Greece, San Marino, Slovakia and Slovenia.

The text of the convention was agreed by the Council of Europe after several years of debate. It sets out major principles and rules that signatory states must incorporate into domestic law, including bans on

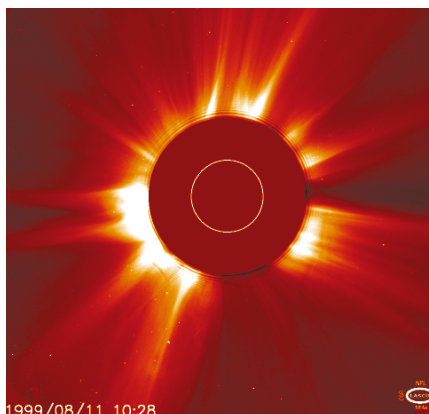
discrimination on the grounds of a person's genetic make-up and predictive genetic tests for other than medical purposes. It also prohibits the creation of human embryos for research purposes, and financial gain from the use of any part of the human body.

Former lab boss faces the rap over spy scandal

[WASHINGTON] Bill Richardson, the US energy secretary, has recommended that unspecified disciplinary action be taken against Sig Hecker, the former director of the Los Alamos National Laboratory in New Mexico, for his role in developments that led to the Chinese spying scandal at the laboratory.

An investigation by the Department of Energy's inspector general found that an official, subsequently identified as Hecker, failed to implement the department's request to limit the computer access of Wen Ho Lee, the Los Alamos scientist suspected of leaking nuclear weapons design information to China.

John Browne, who succeeded Hecker as director of Los Alamos in 1997, must now decide what action to take against Hecker, who still works at the laboratory as a researcher.



Lights out: the sun seen last week from the SOHO spacecraft (left) and from Earth during an eclipse.

Eclipse shows sun's corona in a fresh light

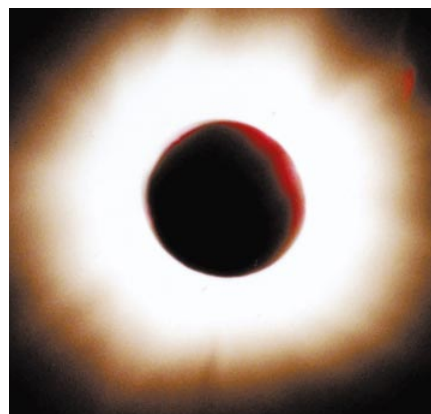
These images compare the Sun's corona as seen 'virtually' (left) by the LASCO C2 coronagraph on the SOHO spacecraft, a joint mission by the European Space Agency and the US space agency NASA, and from the ground, both taken during last week's full solar eclipse as seen in Northern Europe.

The latter image was taken with a 500 mm lens in the town of Le Treport in northern France. Both show solar north at the top, west on the right, and the corona near the time of totality. Rotated through 45 degrees to

compare it with the SOHO picture, the ground-based photograph shows the corona extending around three or four solar radii.

There were significant changes in the coronal structure in the few days before the eclipse. In particular, several large coronal mass ejections were observed by the LASCO telescopes. In both images, streamers can be seen coming from high-latitude regions.

The bright region to the right just above the equator is the remnant of a coronal mass ejection which erupted early on 7 August. The region just below the equator on the left is the remnant of another mass ejection which occurred one day before the eclipse.



Bureaucrats pose threat to museums

Sir—Your article on the funding difficulties of Seattle's Burke museum, and its bid for independence, illustrates a common problem faced by university natural history museums (*Nature* 399, 189; 1999). The article also mentioned the 100-year-old Sam Noble Oklahoma Museum of Natural History—the largest repository of Oklahoma's heritage. As I was offered the directorship of the Burke and have been director of the Sam Noble museum for 17 years, my views may be germane.

Today, university presidents are selected from a rotating pool of administrators who no longer have a deep understanding of a particular campus. They do not possess any special institutional loyalty. If they avoid offending the governing board, raise enrolment, and increase income, they will succeed. Soon they will move to a more prestigious position. A museum's future hinges on decisions made by these transient administrators, but what can heritage mean to such a person, and how can such a transitory bureaucrat be expected to deal

with the special needs of an enduring institution?

The intangible value of scholarship by museum curators, educational benefits for students, and public outreach are small change for administrators focused on costs and benefits. Their concerns are the high cost of storing collections, the low funding for museum research, the decline of taxon-based classes that require specimens, and the low public profile of campus museums. Does the museum produce credit hours to justify more funds? Does it make money? Does it have public support that could cause political problems? The answer to each is usually 'No'.

Nevertheless, one hears remarkable success stories. The University of Oklahoma has built an outstanding \$44 million building (www.snomnh.ou.edu), although Oklahoma is a poor state. We, too, were told that a new museum was impossible, but we took the project to the people and drew on their power for support.

University museums, as keepers of the

nation's heritage, must develop a long-term strategy that calls on their constituents to understand and support their purposes. It is dangerous for a museum to pull away from its university. Will a private board understand the importance of collections when not used for teaching and research? Will it only be interested in exhibits? Will it see the need for curators?

University museums provide leading scholars in disciplines from systematics to conservation. They train the scientists of tomorrow. Museums must not lose sight of their unique role—to preserve and interpret heritage, contribute knowledge, and educate students and the public. Leaving the university violates the museum's *raison d'être*. The present administrators will soon be gone, but the people—the ultimate supporters—will remain. Museums are forever.

Michael A. Mares

*Sam Noble Oklahoma Museum of Natural History,
University of Oklahoma,
Norman, Oklahoma 73072, USA*

What price 'prestige' in publishing?

Sir—I support Mark Riley in his protest against the outrageous price increases of some journals, but I do not think that all the blame rests with the commercial publishers (*Nature* 399, 623; 1999). They are in the business of making profits. We should ask ourselves why the scientific community lets these people run our research journals to their, rather than our, advantage.

The answer is in the fallacious system of 'journal prestige' which scientists have created. Nothing stops us from communicating our research results through less costly systems, such as preprints, electronic self-publishing, and small-scale journals run by interest groups.

But, unless a paper is published in a 'prestigious' journal, it does not count for much for grants, promotions, invited talks, and so on. This is how we let commercial publishers keep a grip on our throats. It is within our capacity to change it, if we wish. And, unless we do so, the commercial publishers cannot be blamed if they run our business as they see fit.

Although perceptual changes will be difficult to achieve by legislation from above, some much-needed adjustment in the research community can be started at the grassroots level. For example, any 'official' use of journal prestige ranking

should be resisted or drastically minimized. Our obsession with numerical ranking of almost everything has gone too far and often brings more harm than good.

Second, the research funding system should encourage publication in local journals. In Canada, for example, it is considered almost a disgrace to publish in the 'Canadian Journal of XYZ' in some disciplines. Pressure to publish in only some celebrated journals has become an unhealthy obsession fuelled by the reward system in academic institutions.

Finally, the idea of discount publishing at the lowest possible commercial rates should be considered.

Alexander A. Berezin

Department of Engineering Physics, McMaster University, Hamilton, Ontario, Canada L8S 4L7

Icelanders opt out of genetic database

Sir—In a News article¹ the ethical guidelines for a Swedish biobank were compared to the lack of ethical considerations in the Health Sector Database (HSD) law in Iceland. The Icelandic project, proposed by deCODE Genetics, Inc, fares poorly in the comparison, mostly because the law permits the granting of an exclusive licence to a single company to do genetic and other

research on the whole-population database, without requiring informed consent. Instead the term presumed consent is introduced, leaving an opting-out possibility for those who fill in the correct forms.

Jeffrey Gulcher and Kari Stefansson of deCODE object to the comparison². Some inaccuracies in their letter deserve correction. First, they state that, according to the HSD law, medical information can be linked to genealogy and genetic information only with informed consent. But informed consent is not mentioned in the law³. No ethics committee will have the authority to require consent, unlike in the Swedish project.

Second, they state that "the bill grants deCODE the exclusive right to market this anonymous database outside Iceland, but not the exclusive right to use the database". This is also incorrect. Although the HSD law was sponsored by deCODE the company is not mentioned in it, and the licence has not been granted yet. More importantly, the database will not be anonymous since it will contain personal identifiers which, although encoded, can be used to identify individuals⁴.

Third, the law grants the licensee the exclusive rights to create and operate the database, and says nothing about the access of other researchers, contrary to what Gulcher and Stefansson claim.

Fourth, the authors say that "deCODE obtained its licence to construct and run the

health-care database through the democratic process". But important changes were made to the bill in parliament, without debate. For example, the bill now allows the licensee to connect the medical database to a genetic database, without requiring informed consent. The warnings of nearly all the independent expert groups that were asked to comment on the bill were ignored³. Our small society was not able to withstand deCODE's expensive information campaign.

Now the state-controlled banks have purchased almost half of the US venture capitalists' original investment in deCODE, increasing concern about the close ties between deCODE and the government.

The European Union's Data Protection Commissioners recommended that the Icelandic authorities should reconsider the project in the light of the European Convention on Human Rights. This was not done, and the consequences are lack of consent, lack of traditional ethics control and lack of freedom to withdraw information entered into the database.

More than 11,000 Icelanders have opted out of the HSD database. Many doctors have promised not to send information about patients to the database, so we believe it will not be created as originally envisaged. Sweden's UmanGenomics seems to be doing a much better job.

Pétur Hauksdóttir

Mannvernd, Association for Ethics in Medicine and Science, PO Box 94, Reykjavik 112, Iceland

Sir — Gulcher and Stefansson's letter² is a striking demonstration of the corporate culture of companies such as deCODE Genetics and British Biotech, and their difficulty in telling the story straight⁵.

To assert that the HSD law was approved through a democratic process underestimates the steamrollering power of a large government majority. Warnings from the Icelandic Medical Association, local and international geneticists, and privacy experts were ignored. The speed of the Icelandic legislative process precluded balanced and informed analysis of a complicated issue. It is sheer spin-doctoring to suggest that this over-speedy, ill-thought-through legislation expresses informed community consent.

Using encrypted identifiers for a comprehensive dynamic database describing the health status, genealogy and genotype of a population without informed consent opens up an ethical Pandora's box.

Skúli Sigurdsson

Institute of Cultural Studies, Humboldt University, Sophienstrasse 22a, 10178 Berlin, Germany

1. Abbott, A. *Nature* **400**, 3 (1999).

2. Gulcher, J. R. & Stefansson, K. *Nature* **400**, 307–308 (1999).

3. <http://www.mannvernd.is/english>

4. Anderson, R. Br. Med. J. **319**, 59 (1999).

5. *Nature Biotechnol.* **16**, 491 (1998).

Corals resist extinction by global warming

Sir — Coral reefs form complex ecosystems that easily reflect changes in environmental conditions. It is therefore alarming that many reefs show signs of bleaching, which can be interpreted as the beginning of degeneration. Peter Pockley reports that experts ascribe the degeneration to global warming, and predict the disappearance of most reefs within a century (*Nature* **400**, 98; 1999). This prediction may be based on biological considerations, but it seems to contradict geological data.

Coral reefs have survived geological periods with considerably higher — and others with considerably lower — temperatures than we face now or in the near future. The reefs also survived sea level fluctuations between more than 10 metres above the present level and more than 100 metres below. So it appears from geological history that corals — like most organisms — are well capable of adapting to changing environments, even though they may be less flourishing during a period of change.

The geological past shows that global warming in itself is not a threat to reefs.

A. J. (Tom) van Loon

Geocom, P. O. Box 336, 6860 AH Oosterbeek, The Netherlands

'Snowball Earth' theory still stands

Sir — George Williams¹ defends the high-obliquity hypothesis for low-latitude glaciation near sea level during the Proterozoic². He criticizes the 'snowball Earth' hypothesis³ on the grounds that global glaciation would be accompanied by drastic lowering of sea level. Sea level is lowered when ocean water is sequestered as land ice, but the volume of land ice that would exist in a 'snowball Earth' is uncertain because the hydrologic cycle would be severely reduced if the oceans were frozen over⁴.

There is clear evidence for emergence and hiatus during the Marinoan glaciation (~600 Myr ago) in Australia⁵, "expressed by an erosional disconformity, in places with large-scale channelling". The incised channels are ~150 metres deep⁶, which would represent a minimum for the lowering of sea level assuming the land was depressed by glacial ice. This exceeds the lowering of sea level accompanying any Phanerozoic glaciation.

Williams¹ points to the presence of seasonal freeze-thaw structures⁷, which he says could not form near the Equator

because of low seasonal temperature variation. In fact, periglacial ice-wedge structures occur at 20°N latitude in Hawaii⁸ and near the Equator on Mount Kilimanjaro in Tanzania⁹. They are attributed to diurnal freeze-thaw cycles and may be shallower than those associated with the Marinoan glaciation⁷. Ice-wedge polygons are less well developed than linear ice wedges in the mountains because of surface slopes.

The 'snowball Earth' theory accounts for other features of the Neoproterozoic sedimentary record³ — banded iron formations, post-glacial cap carbonates, and large carbon-isotopic variations. The high-obliquity hypothesis provides no explanation for these features.

Paul F. Hoffman

Department of Earth Sciences, Harvard University, Cambridge, Massachusetts 02138, USA

1. Williams, G. E. *Nature* **398**, 555–556 (1999).

2. Williams, G. E. *Earth Sci. Rev.* **34**, 1–45 (1993).

3. Hoffman, P. F. *et al.* *Science* **281**, 1342–1346 (1998).

4. Hoffman, P. F. & Schrag, D. P. *Science* **284**, 1087a (1999).

5. Preiss, W. V. *Geol. Surv. S. Aust. Bull.* **53**, 438 (1987).

6. Dyson, I. A. & von der Borch, C. C. in *Incised-valley Systems: Origin and Sedimentary Sequences* (eds Dalrymple, R. W. *et al.*) 209–222 (Soc. Sediment. Geol. Spec. Publ. 51, Tulsa, OK, 1994).

7. Williams, G. E. *Precambrian Res.* **32**, 233–242 (1986).

8. Gregory, H. E. & Wentworth, C. K. *Bull. Geol. Soc. Am.* **48**, 1719–1742 (1937).

9. Downie, C. & Wilkinson, P. *The Geology of Kilimanjaro* (Dept Geol., Univ. Sheffield, Sheffield, 1972).

Visit heaven and hell ahead of schedule

Sir — It is surprising to read that creationists in Kansas are lobbying the state board of education to include their views in school science teaching¹. A literal reading of the Bible can sometimes be absurd, and it contains many numerical data that reflect only the views of an earlier age. These numbers can be the origin of funny calculations². A colleague and I published a squib in which, using paragraphs from the biblical books of Isaiah and Revelation, the temperatures of heaven and hell were calculated as 504.5 and 716.6 K respectively³.

The calculation was discussed widely in the media. In the UK *Sunday Times* (9 August 1998), a geophysicist suggested possible locations, based on our results. Hell might be the hydrothermal vents on the bottom of the ocean, and heaven could be the thermosphere. It must be satisfying for creationists that such places can now be reached, thanks to the work of scientists.

Jorge Mira-Pérez

Departamento de Física Aplicada, Universidade de Santiago de Compostela, E-15706 Santiago de Compostela, Spain

1. Dalton, R. *Nature* **398**, 453 (1999).

2. Peil, K. *Nature* **399**, 522 (1999).

3. Mira-Pérez, J. & Vina, J. *Physics Today* **51** (7), 96 (1998).

Neptune's misbehaving rings

Mark R. Showalter

Without confinement, the arcs of material orbiting Neptune would have spread out into rings long ago. But new observations show they are still there, challenging astronomers' best attempts at explaining them.

Consider this thought experiment. You are orbiting a planet and you open up a big bag of marbles. What happens? Well, at first the marbles stay together, because they and you are following roughly the same orbital path. But soon, jostling among the marbles will occur as they bounce off one another. Eventually they will spread apart, with each marble following a slightly different path. Some marbles will follow slightly lower orbits that circle the planet faster, whereas others will follow higher orbits that circle the planet more slowly. Eventually, the faster-moving marbles will lap the slower-moving ones, and you will find yourself within a continuous, uniform ring encircling the entire planet.

What could you do to prevent the marbles from spreading? Well, it wouldn't be easy. This is why the discovery in 1984 of an incomplete ring of material orbiting Neptune¹ represented such a dynamical puzzle — an arc shouldn't last for more than a few months. On pages 731 and 733 of this issue^{2,3}, two groups of astronomers report that the ring arcs of Neptune are alive and well, and still narrowly confined in space nearly 15 years after their discovery. Furthermore, they find that the arcs' current orbital position invalidates the most promising model of their confinement; for dynamicists, it is time to go back to the drawing-board.

Twenty-three years ago, Saturn was the only planet in our Solar System known to have rings. In the images of the day, Saturn's rings seemed to confirm dynamicists' crude expectation that all rings should be simple — broad, featureless, flat, equatorial, circular and symmetric — because collisions between ring particles would rapidly destroy any deviations from this set of properties. Since then, the pillars of this conventional wisdom have fallen one by one, with observers in the lead and dynamicists struggling to keep up.

Back in 1977, astronomers were observing Uranus as it passed in front of a particularly bright star. Such stellar 'occultations' make it possible to probe the atmospheres of distant planets with great sensitivity. On this particular occasion, Elliot *et al.*⁴ noticed that the star blinked out briefly five times before the planetary occultation, and again five times afterwards. They had discovered the

first five of Uranus' ten narrow rings. Unlike Saturn's broad rings, narrow rings require a confinement mechanism to prevent them from spreading radially. This led to the 'shepherding' model of ring confinement, in which gravitational perturbations from nearby moons are thought to define abrupt ring boundaries⁵.

Hubbard *et al.*¹ were using the same occultation technique when they discovered the ring arcs of Neptune. This time the star blinked out on only one side of the planet, implying an incomplete ring. From this surprising result, two models for arc confinement emerged. Goldreich *et al.*⁶ showed that, if the ring and a nearby moon had a particular type of orbital relationship called a co-rotation resonance, the moon could drive ring material into a set of uniformly spaced 'libration sites', with the material unable to cross between sites. The model seemed to predict a 'beaded ring' of regularly spaced clumps, and it placed very specific constraints on the relationship between the co-rotation resonance and the number of beads. In an alternative shepherding model, Lissauer⁷ showed that perturbations by a moon embedded directly in the ring could stabilize arcs 60° (or one-sixth of the orbit) away, if another shepherding moon was also nearby.

It was in this context that Voyager 2 arrived at Neptune in 1989, finally showing the world what incomplete rings look like (Fig. 1). Images showed a faint ring containing a few brighter arcs all grouped together within a 40° sector. At first glance, it appeared incompatible with both models. Fortunately, the Voyager images also revealed a nearby moon, Galatea, orbiting just inside the ring. By carefully checking the orbital motions, Porco⁸ showed that Galatea was indeed in, or at least very close to, a co-rotation resonance that should generate 86 stable libration sites in the ring, each 4.2° long. She argued that the Goldreich model could explain the data, providing only a subset of these sites happened to be filled with material. This model was almost universally accepted, because it would otherwise be too much of a coincidence for the arcs to fall so close to a co-rotation resonance.

Nonetheless, there were a few minor problems. First, the observed width of the ring was too great to fit neatly inside the (very



Figure 1 This classic Voyager image shows two of Neptune's faint rings. The outer ring contains three bright clumps, which are the most prominent ring arcs, known as Liberté, Egalité and Fraternité. A fainter, fourth arc, Courage, falls outside this field of view. The inner LeVerrier ring is also visible. Nine years after this image was taken, new observations from Earth² and from space³ conflict with the predictions of the most promising model for arc confinement, reopening the debate about the arcs' dynamics.

narrow) co-rotation resonance. Second, the finest Voyager images clearly showed very small clumps of material within individual arcs, yet the model was unable to explain features less than 4.2° apart.

In the ten years since the Voyager encounter, observing techniques have improved markedly. Now, for the first time, the arcs of Neptune have been imaged from the ground² and from a telescope in orbit around the Earth³. The advantage of these new observations is that they show global views of the arcs, rather than the single 'slice' of the system obtained by occultation studies. Unhappily, the new detections place the arcs about 20° away from the Goldreich model's prediction and definitely outside the co-rotation resonance^{2,3}. This surprising result will force dynamicists to rethink once again the dynamics of arc confinement. The current consensus seems to support a variant on the Lissauer model⁷, in which a small number of unseen moons confine arcs between them⁹. Unfortunately, such a model has enough free parameters as to be virtually impossible to test at this time. The moons needed to confine the arcs could be quite tiny, say 10–20 km in diameter, placing them far beyond the capabilities of Earth-based observations. We may have to wait until another spacecraft visits Neptune to settle this question.

Models of planetary ring systems have thrown up a recurring puzzle — although



100 YEARS AGO

The action of magnetism on the propagation of light in a transparent medium has been rightly regarded as one of the most beautiful of Faraday's great scientific discoveries. Like most important discoveries it was not result of accidental observation, but was the outcome of long and patient inquiry. Guided by a conviction that (to quote his own words) "the various forms under which forces of matter are made manifest have one common origin," he made many attempts to discover a relation between light and electricity, but for very long with negative results. Still, however, retaining a strong persuasion that his view was correct, and that some such relation must exist, he was undiscouraged, and only proceeded to search for it more strictly and carefully than ever. At last, as he himself says, he "succeeded in magnetising and electrifying a ray of light, and in illuminating a magnetic line of force."

From *Nature* 17 August 1899.

50 YEARS AGO

Osborne Reynolds, in a short paper entitled "On the Action of Rain to Calm the Sea", discussed the action of rain in calming a rough sea. He expressed the opinion that each drop of rain produces a vortex ring which, on descending into the water, transfers momentum from the surface layers to the underneath layers, thus reducing the relative motion of the layers. It was suggested to us by Prof. A. H. Gibson that further experiments were desirable to confirm the above theory. It has now been found that whereas vortex rings are, in fact, produced when the height of dropping, and hence the velocity of the drop, is small, only 'splash and surface effects' are produced when the height is great. ... It may be concluded, therefore, that very few, if any, of the drops in a normal rain-storm actually produce vortex rings when they strike the surface of water. Rather is the kinetic energy of the drop dissipated in shock and splash at the surface, and thus the sphere of influence of the drop is not as great as suggested by Reynolds.

From *Nature* 20 August 1949.

rings now appear to be common features of the giant planets, their expected lifetimes are far shorter than the age of the Solar System. Rings should eventually collapse because of gravitational interactions with nearby moons, and the moons themselves should be destroyed by meteoroid impacts. The emerging view is that the ring-moon systems are not ancient structures existing since the Solar System began, but transient phenomena in which the death of a moon is also the birth of a new ring. If this is so, then the stable ring arcs of Neptune are of considerable interest as an intermediate structure between accreted moons and distributed rings. Perhaps we are seeing the cloud of debris from a disrupted moon, whose expansion was arrested by its chance proximity to a family of resonances capable of confining it.

Despite these puzzles, rings remain useful as dynamical laboratories in which astronomers can observe the processes that work in much larger astrophysical systems, such as galaxies and protoplanetary disks. On one level, the new reports are noteworthy simply as illustrations of how far astronomical methods have improved in the past decade, enabling two teams to detect a tiny feature a

fraction away (less than two arc seconds) from a planet that is 10^6 times brighter. The ability to image such features directly will reduce the need for astronomers to wait patiently for the occasional bright star to pass behind the rings (although stellar occultations will always remain useful because of the extraordinary geometric precision they provide). Even though the interactions between Neptune's moons and ring arcs just became a great deal murkier, the good news is that the dynamics of some of the Solar System's oddest structures are becoming much easier to study. □

Mark R. Showalter is in the Space

Telecommunications and Radioscience Laboratory, Stanford University, Stanford, California 94305, USA. e-mail: showalter@ringside.arc.nasa.gov

1. Hubbard, W. B. *et al.* *Nature* **319**, 636–640 (1986).
2. Sicardy, B. *et al.* *Nature* **400**, 731–733 (1999).
3. Dumas, C., Terrile, R. J., Smith, B. A., Schneider, G. & Becklin, E. E. *Nature* **400**, 733–735 (1999).
4. Elliot, J. L., Dunham, E. W. & Mink, D. J. *Nature* **267**, 328–330 (1977).
5. Goldreich, P. & Tremaine, S. *Nature* **277**, 97–99 (1979).
6. Goldreich, P., Tremaine, S. & Borderies, N. *Astron. J.* **92**, 490–494 (1986).
7. Lissauer, J. J. *Nature* **318**, 544–545 (1985).
8. Porco, C. C. *Science* **253**, 995–1001 (1991).
9. Sicardy, B. & Lissauer, J. J. *Adv. Space Res.* **12**, 1181–1195 (1992).

Phytochromes

Tripping the light fantastic

Harry Smith

Photosynthesis fuels life as we know it, so we should all be grateful that evolution has provided plants with sensory mechanisms that allow them to adapt precisely to sunlight. These signal-transducing photoreceptors tell the plant about the intensity, spectral quality, direction and periodic timing of solar radiation, setting off metabolic and developmental changes that optimize the photosynthetic absorption of photons. The most intensively researched photoreceptors are the five members of the phytochrome family (known as phytochromes A to E in the tiny weed *Arabidopsis*). Phytochromes are photochromic pigments — that is, they are photochemically converted from the biologically inactive Pr form to the active Pfr form by absorbing red photons. Conversely, Pfr can be converted back to Pr by absorbing far-red photons.

For more than 40 years, the Holy Grail of phytochrome research has been the molecular mechanism of Pfr action. Ni *et al.*¹ now provide compelling data (on page 781 of this issue) about an interaction between phytochrome B and its putative reaction partner, phytochrome-interacting factor-3 (PIF3). The authors also speculate on a short-cut mechanism that links phytochrome photo-conversion to gene regulation. This work is the latest stage in a sustained attack on the mechanism of phytochrome action from the

laboratory of Peter Quail. And it is the first demonstration of photoreversible binding of a phytochrome to its putative signalling partner *in vitro*.

Quail's laboratory had already identified PIF3 by two separate approaches. In the first, Ni *et al.*² used a yeast two-hybrid screen to fish for factors that interact with non-photoactive carboxy-terminal fragments of phytochrome A and phytochrome B. (Phytochrome molecules are thought to act, at least in part, through a region in the carboxy terminus.) Ni and colleagues found² that PIF3 binds to native phytochrome A and phytochrome B. Overexpression of PIF3 in the sense orientation increased light sensitivity; overexpression in the antisense orientation strongly decreased it. Crucially, binding of PIF3 was reduced with a series of phytochrome A and phytochrome B molecules carrying missense mutations that caused loss of *in vivo* activity. In the second approach, Halliday *et al.*³ isolated a new mutant of *Arabidopsis* with enhanced sensitivity to red light. The mutation turned out to be an insertion in the promoter of the *PIF3* gene, causing the gene to be overexpressed. This study provided further evidence that PIF3 is involved *in vivo*, and confirmed that it positively regulates phytochrome action.

Perhaps the most intriguing features of PIF3 are its nuclear localization and

characterization as a basic helix–loop–helix protein with putative transcription-factor activity. Phytochrome action has long been thought to involve the selective regulation of gene expression, and there is evidence that the phytochrome translocates into the nucleus after conversion to Pfr⁴. Ni *et al.*¹ now show the photoreversible interaction of phytochrome B with PIF3 *in vitro*. The dynamic characteristics of this interaction fit the photobiological *in vivo* information. The authors show that PIF3 binds to phytochrome B only after the phytochrome has been photoconverted from Pr_B to Pfr_B. Reconversion of Pfr_B to Pr_B by far-red light abruptly terminates the association. The binding of PIF3 to Pfr_B seems rather slow compared with *in vivo* response kinetics. However, the binding was done at 4 °C, and dissociation on treatment with far-red light was reassuringly rapid.

As in previous experiments, missense mutations in phytochrome B caused sub-

stantial reductions in binding to PIF3. Although details of the binding process still need to be worked out (for example, stoichiometry data are not yet available), binding sites in both the amino- and carboxy-terminal domains are essential for full association. We already know that photoisomerization of the light-absorbing moiety (chromophore) causes conformational changes in the holoprotein⁵. Now, the data of Ni *et al.*¹ support a simple hypothesis involving three steps: translocation of Pfr_B to the nucleus; binding of Pfr_B to the constitutively nuclear PIF3; and interaction of the Pfr_B–PIF3 complex — either directly or indirectly — with target genes.

The apparent simplicity of this idea may, of course, turn out to be illusory. Other factors bind to phytochromes (including the cryptochrome photoreceptor CRY1)⁶. Indeed, one such factor (phytochrome kinase substrate-1, PKS1), fished out by the two-hybrid screen, is a substrate for

phytochrome A-mediated serine/threonine kinase activity⁷. There has been much controversy about whether the phytochromes are kinases⁸. But the accumulated evidence, including homologies with phytochrome-like molecules in the cyanobacteria^{9,10}, is so strong that the question now is the extent to which their kinase properties are implicated in signal transduction.

Earlier this year Fankhauser *et al.*⁷ showed that the cytoplasmically located PKS1 binds to, and is phosphorylated by, both Pr_A and Pfr_A, but that phosphorylation activity is between 2- and 2.5-fold higher with the Pfr form. If this were to initiate a kinase cascade, it could be enough to explain the huge signal amplification observed *in vivo*. But these authors propose an intriguing alternative. Based on analogy with activation of NF- κ B in mammalian cells¹¹, they suggest that binding to PKS1 may be important for retaining phytochrome A in the cytoplasm. In this way, translocation of the phytochrome to the nucleus would be negatively regulated. Certainly, PKS1 negatively regulates the action of phytochrome A, as judged from overexpression experiments⁷. So, PKS1 may be a cytoplasmic anchoring protein, and phytochrome-mediated phosphorylation of PKS1 might release phytochrome A to move into the nucleus.

The differences between the results of Ni *et al.*¹ and Fankhauser *et al.*⁷ may lie in their choice of phytochrome A or phytochrome B as the experimental material, although PIF3 and PKS1 bind to both. It is premature to try to resolve all the questions about phytochrome action, and Fig. 1 shows only one of several possibilities. Such a scheme could account for phytochrome-mediated responses that operate through gene regulation, or through other actions — rapid alterations in growth rate or membrane properties, or chloroplast relocation — that probably emanate from a separate primary action located in the cytoplasm.

The transition between molecular genetics and biochemistry is never easy, and major caveats need to be mentioned. For example, mutants provide evidence for both common and separate pathways of phytochrome A and phytochrome B action. Furthermore, pharmacological evidence points towards a number of intermediary metabolites (such as guanine-nucleotide-binding (G) proteins, Ca²⁺ and cyclic GMP) in signal transduction. Integrating all these potential branches into a single scheme will test our ingenuity. The first steps in the choreography of phytochrome signal transduction are beginning to be worked out — but will the dance be a *pas de deux* or a hoedown? □

Harry Smith is in the Department of Biology, University of Leicester, Leicester LE1 7RH, UK.
e-mail: has@leicester.ac.uk

1. Ni, M., Tepperman, J. M. & Quail, P. H. *Nature* **400**, 781–784 (1999).

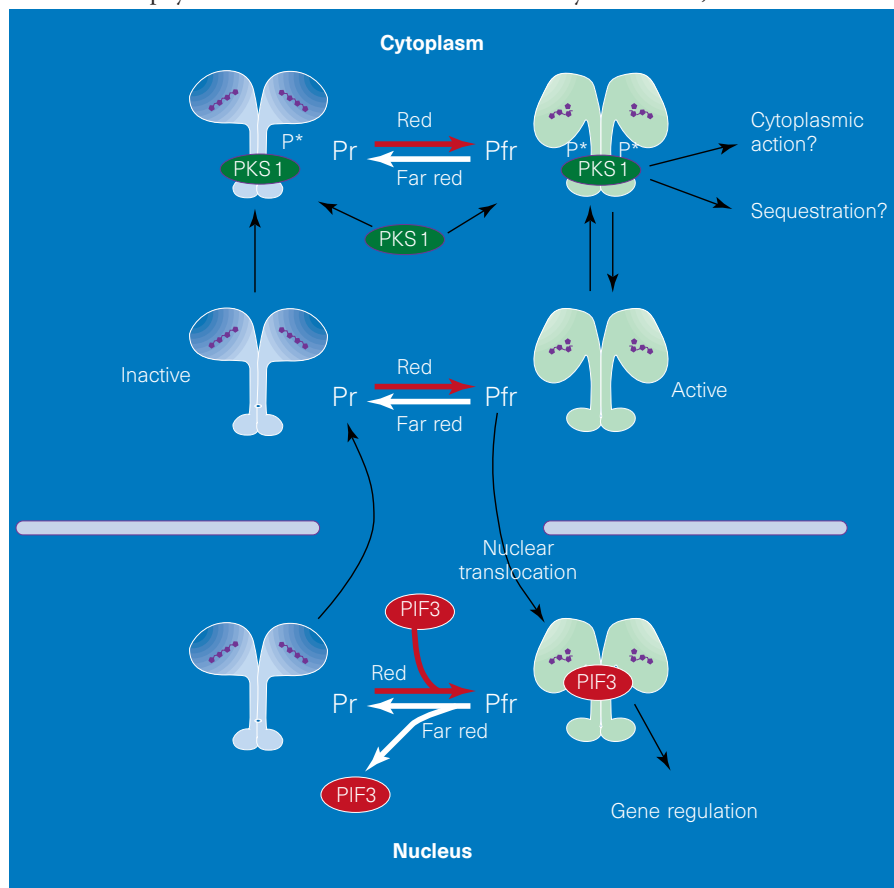


Figure 1 Main steps in phytochrome action. Absorption of a red photon by the inactive Pr form of a phytochrome causes a conformational change in the dimeric photoreceptor molecule. In the Pfr form, the phytochrome translocates to the nucleus where it binds to a putative reaction partner, PIF3, which is constitutively found in the nucleus and has the characteristics of a basic helix–loop–helix transcription factor. The Pfr–PIF3 complex initiates gene regulation, either directly or through unknown intermediates. Reversion of Pfr to Pr by far-red light results in rapid dissociation of PIF3, interrupting signal transduction. In the Pr form, the phytochrome slowly relocates to the cytoplasm. Here, in either the Pr or Pfr form, it can bind to a kinase substrate (PKS1), which may be involved in retention of the phytochrome in the cytoplasm or in its release for translocation. Phosphorylation of PKS1 is enhanced with Pfr, and may be the prelude to cytoplasmic action. So, several steps are susceptible to regulation by absorption of light by the photoreceptor: phosphorylation, nuclear translocation, association with PIF3 and transfer of signal transduction to PIF3.

2. Ni, M., Tepperman, J. M. & Quail, P. H. *Cell* **95**, 657–667 (1998).
3. Halliday, K. J., Hudson, M., Ni, M., Qin, M. & Quail, P. H. *Proc. Natl Acad. Sci. USA* **96**, 5832–5837 (1999).
4. Yamaguchi, R., Nakamura, M., Mochizuki, N., Kay, S. A. & Nagatani, A. *J. Cell Biol.* **145**, 437–445 (1999).
5. Song, P.-S. *J. Biochem. Mol. Biol.* **32**, 215–225 (1999).
6. Ahmad, M., Jarillo, J. A., Smirnova, O. & Cashmore, A. R. *Mol. Cell* **1**, 939–948 (1998).
7. Fankhauser, C. *et al. Science* **284**, 1539–1541 (1999).
8. Cashmore, A. R. *Proc. Natl Acad. Sci. USA* **95**, 13358–13360 (1998).
9. Kehoe, D. M. & Grossman, A. R. *Science* **273**, 1409–1412 (1996).
10. Yeh, K. C., Wu, S. H., Murphy, J. P. & Lagarias, J. C. *Science* **277**, 1505–1508 (1997).
11. Stancovski, I. & Baltimore, D. *Cell* **91**, 299–302 (1997).

Atmospheric chemistry

Sulphur emissions from ships

Barry J. Huebert

A decade ago, the idea that oceanic phytoplankton regulate climate¹ by producing dimethyl sulphide (DMS) led to the conventional wisdom that only natural DMS-derived sulphate aerosols influence cloud reflectivity (and thus the balance of radiation reaching the Earth) in remote marine areas. According to Capaldo *et al.*² (on page 743 of this issue), there is another important sulphur source in the post-industrial era: sulphur dioxide pollution from ships burning fossil fuels. The global atmospheric sulphur model used by Capaldo *et al.*, which is the first to incorporate a new inventory of emissions from worldwide shipping³, not only tells us that the effect of human activity on aerosols and climate is greater than we previously thought, but also that attempts to fit earlier models to data may have been misleading. This isn't the first time air pollution has been poorly quantified, but it is worth investigating how the omission of ship exhausts may have led to misunderstandings in other aspects of atmospheric chemistry models.

Clouds reflect more sunlight back to space when they form on larger numbers of aerosol particles. It has been argued¹ that more DMS might generate more sulphate aerosols and therefore create greater concentrations of cloud-condensation nuclei, which would make clouds brighter and cool the Earth a little (Fig. 1). Now we learn that shipping may generate more SO₂ above parts of the oceans in the Northern Hemisphere than DMS does. The Capaldo *et al.* study² is revealing in what it says about the confidence we should place in numerical models. Previous models neglected ships, but now that we have a reasonable inventory of ship emissions, the predictions are naturally quite different. This work represents a substantial change in our picture of atmospheric sulphur over much of the Earth. I think it's good for us to realize now and then that we haven't discovered everything of interest, and surprises are still out there.

We gain confidence in models as representations of reality when they predict spatial and temporal patterns that we can observe independently. Measurements — on the scale of the phenomena the models are describing — are the sole means of

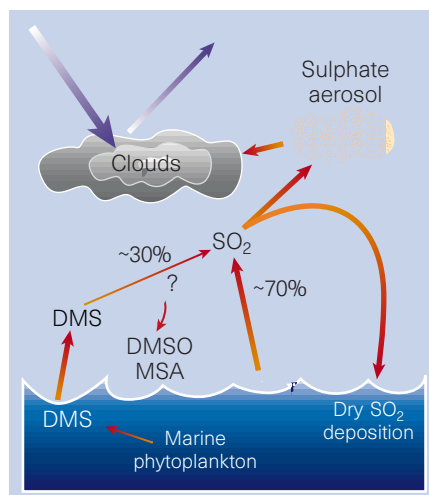


Figure 1 Sources and sinks of sulphur in the middle of the North Atlantic. Capaldo *et al.*² show that ship emissions generate far more SO₂ than previously assumed where there is heavy ship traffic. The main natural source of SO₂ over the oceans comes from oxidation of dimethyl sulphide (DMS) produced by phytoplankton. Some of this SO₂ will be oxidized to sulphate aerosols that affect the reflectivity of clouds, thereby cooling the Earth. Previous climate models did not consider ship emissions, and so had to derive all their SO₂ from DMS, using estimates of net yields (? mols of SO₂ produced per mol DMS consumed) of almost 1. Models with smaller net yields may now do a better job of matching observations, because much of the observed SO₂ may be derived from ships. DMSO, dimethyl sulphoxide; MSA, methanesulphonic acid.

testing model predictions. Capaldo *et al.*² make an effort to compare their model results with measured SO₂ concentrations, but the results are not very satisfying. The model overpredicts most concentrations of SO₂; the addition of ship emissions causes only a slight improvement. This may be due in part to the fact that we have very few remote SO₂ measurements with which to constrain sulphur models: the need for high-quality time-series data on sulphur species at remote sites is acute. Clearly, we still have a long way to go.

Nonetheless, this new inventory may change some of the parameters used in mod-

els of the pre- and post-industrial atmospheres. A good example is the branching of DMS products between SO₂ (some of which might form sulphate particles) and a variety of other species, including dimethyl sulphoxide and methanesulphonic acid (Fig. 1). The net yield of SO₂ is defined as the fraction of SO₂ produced from oxidized DMS. Published net yields of SO₂ from DMS^{4–6} range from 0.1 to almost 1, so there is a lot of room for improvement. In fact, even models of pre-industrial sulphur chemistry may be changed by the new data on SO₂ emissions from shipping, because observations of the present (polluted) atmosphere have been used to infer (possibly erroneous) net yields used in pre-industrial models.

It's not an easy matter to determine the net yield of SO₂: laboratory studies simply cannot mimic the wide range of conditions that ambient air experiences. To determine the net yield in the field, one must quantify not only the DMS source, but also the formation and loss rates of all related species. This has been achieved only in extremely clean environments, such as Antarctica⁷ and Christmas Island in the mid-Pacific⁸, where the system is sufficiently simple for one to constrain the important variables. (It remains to be seen, though, whether such pristine environments are good analogues of polluted ones.)

The alternative is to use the net yield of SO₂ as an adjustable parameter⁹: a knob that can be twiddled to make large-scale models fit observations of sulphur species. But, because poorly constrained loss processes and other SO₂ sources (such as the shipping described by Capaldo *et al.*²) also affect the SO₂ concentration, this approach is probably misleading. With the extra SO₂ from shipping, modellers won't have to create all their SO₂ from DMS, so a smaller net yield from DMS might produce sensible concentrations. Clearly, a lot of data will need to be reinterpreted in light of our new knowledge of sulphur emissions from ships.

We now have strong evidence that the 'unpolluted' part of the Earth is considerably smaller than we believed just a year ago. One wonders whether the impact of this newly quantified SO₂ pollution extends beyond the observed reflective tracks that follow shipping routes in satellite images of clouds¹⁰. It is also evident that sulphur emissions from ships can equal or exceed land-based sources in coastal regions, affecting air quality in these areas. The work by Capaldo *et al.* provides the first good argument for reducing sulphur pollution from shipping.

It's not very comforting to know that our waste is affecting an even larger region. It is comforting, though, to see that the tools for making these assessments are getting sharper. □

Barry J. Huebert is in the Department of

Oceanography, University of Hawaii, Honolulu, Hawaii 96822, USA.

e-mail: huebert@soest.hawaii.edu

1. Charlson, R. J., Lovelock, J. E., Andreae, M. O. & Warren, S. G. *Nature* **326**, 655–661 (1987).
2. Capaldo, K., Corbett, J. J., Kasibhatla, P., Fischbeck, P. & Pandis, S. N. *Nature* **400**, 743–746 (1999).
3. Corbett, J. J., Fischbeck, P. S. & Pandis, S. N. *J. Geophys. Res.* **104**, 3457–3470 (1999).

4. Pham, M., Mueller, J. F., Brasseur, G. P., Granier, C. & Megie, G. *J. Geophys. Res.* **100**, 26061–26092 (1995).
5. DeBruyn, W. J., Bates, T. S., Caaney, J. M. & Saltzman, E. S. *J. Geophys. Res.* **103**, 16703–16711 (1998).
6. Saltelli, A. & Hjorth, J. *J. Atmos. Chem.* **21**, 187–221 (1995).
7. Davis, D. et al. *J. Geophys. Res.* **103**, 1657–1678 (1998).
8. Bandy, A. R. et al. *Geophys. Res. Lett.* **23**, 741–744 (1996).
9. Chin, M. et al. *J. Geophys. Res.* **101**, 18667–18690 (1996).
10. Coakley, J. A., Bernstein, R. L. & Durkee, P. A. *Science* **237**, 1020–1022 (1987).

Cell biology

Mast-cell heparin demystified

James L. Zehnder and Stephen J. Galli

Although the clotting of blood (thrombosis) can be life-saving when it stops the bleeding associated with extensive trauma, in settings such as heart attack, stroke and pulmonary embolism, it can cause disease or even death¹. For this reason, heparin — which can prevent or arrest clot formation^{2,3} — is one of medicine's most commonly used drugs. Yet its physiological importance has long been obscure.

Humphries *et al.*⁴ and Forsberg *et al.*⁵ now report (on pages 769 and 773 of this issue) that targeted disruption of the gene for glucosaminyl *N*-deacetylase/*N*-sulphotransferase-2 (NDST-2) produces transgenic mice that lack heparin, revealing a physiological function for this molecule. Heparin, it seems, regulates the types and amounts of other biologically active mediators in the cytoplasmic granules of a group of cells found in almost all tissues^{6,7}. These cells — the mast cells — are effector cells of inflammation and immunity. So, by regulating the mediators in mast cells, heparin may influence the expression of inflammatory and immunological responses.

This is not the physiological function that early work on heparin would have suggested. Heparin was discovered in 1916 by a medical student called Jay McClean, who was helping William Howell to identify a pro-coagulant in tissue extracts². Unexpectedly, McClean discovered an anticoagulant in liver extracts⁸, and Howell named this heparin (after its hepatic origin)⁹. In 1935, Erik Jorpes showed that heparin contains sulphated glycosaminoglycans, and heparin was introduced as an anti-thrombotic therapy soon afterwards^{2,10}. Specific, sulphated pentasaccharide units in heparin glycosaminoglycans (which, in the native heparin proteoglycan, are attached as side chains to a core protein called serglycin) can bind antithrombin III, a protein that circulates in the blood, inducing an allosteric change and greatly increasing its anti-clotting activity^{2,3}.

In 1936, Hjalmar Holmgren discovered that the cytoplasmic secretory granules found in mast cells are an especially rich source of heparin¹⁰. Mast cells are now thought to be the main, if not the only, source of heparin^{2,3}. Receptors on the surface of mast

cells can bind immunoglobulin E antibodies, and binding of these antibodies to antigen leads to aggregation of the receptors. This results in the rapid extracellular release of heparin, histamine, proteases and other mediators from cytoplasmic granules within the mast cells, as well as the release of other mediators not derived from the granules^{6,7}. Mast cells can also be activated independently of immunoglobulin E. Mast-cell activation can lead to bronchoconstriction (in asthma), changes in the calibre and permeability of blood vessels, and other effects in inflammation, immunity and tissue remodelling^{6,7}.

By disrupting the NDST-2 gene in mice, Humphries *et al.*⁴ and Forsberg *et al.*⁵ have now prevented sulphation of heparin glycosaminoglycans, essentially generating mice that lack true (that is, fully sulphated) heparin. The authors found that NDST-2 is not only required for the synthesis of heparin, but also for the formation of normal cytoplasmic granules in certain mast cells^{4,5}.

Disrupting sulphation of glycosaminoglycans of integral membrane proteoglycans can have profound effects on development — for example, the homologue of vertebrate NDST is essential for the Wingless signalling pathway in the fruitfly *Drosophila melanogaster*¹¹. But NDST-2-deficient mice did not seem to have any other obvious developmental abnormalities⁵. Moreover, although neither study describes tests of clotting, the NDST-2-deficient mice may not be much more likely than normal mice to form blood clots under physiological conditions.

This is not as surprising as it might at first seem. In normal animals, heparin does not circulate in the blood but is usually sequestered inside the cytoplasm of 'resting' (non-activated) mast cells² (Fig. 1a). Studies in genetically mast-cell-deficient mice¹² support the hypothesis that the naturally anticoagulant properties of the vasculature reflect the presence of the heparan sulphate proteoglycan — not the heparin proteoglycan — on the surface of the endothelial cells that line blood vessels^{3,12}.

Humphries *et al.* and Forsberg *et al.* also found that certain mast cells in the NDST-2-deficient mice contain cytoplasmic granules with striking morphological and biochemical abnormalities. Both studies indicate that heparin can regulate — apparently by a post-translational mechanism — the levels of certain positively charged protein mediators inside the mast cell. But, depending on the specific mast-cell populations studied, which varied in their maturity, the authors found that NDST-2-deficient mast cells showed minor or no reductions (Humphries *et al.*),

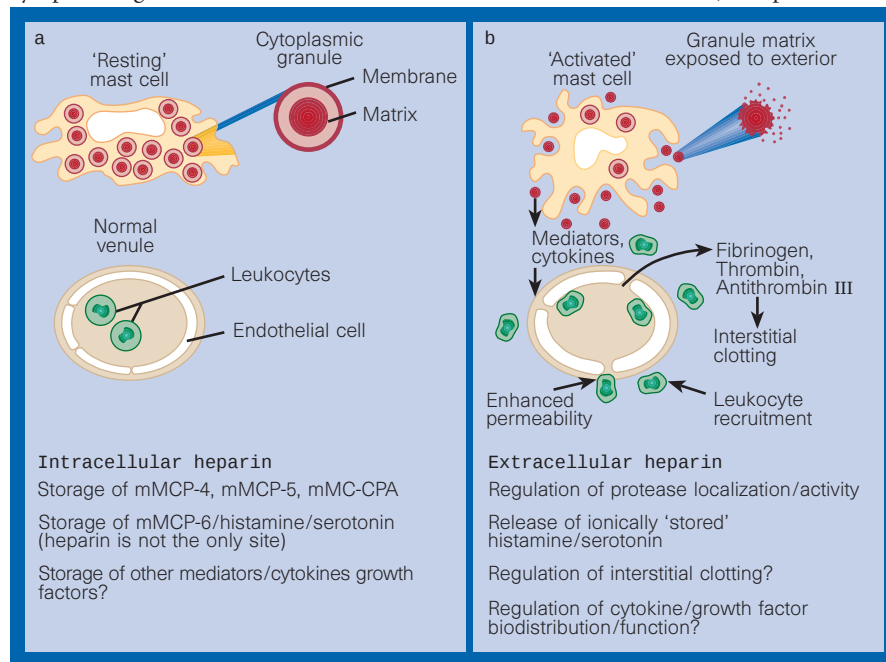


Figure 1 Cell biology and functions of mast-cell heparin, based in part on the findings of Humphries *et al.*⁴ and Forsberg *et al.*⁵. a, Mast cell and venule (small vein) in normal tissue under physiological conditions. The functions and potential functions (?) of heparin stored within the mast cell's cytoplasmic granules are listed. b, An activated ('degranulated') mast cell induces enhanced permeability of the adjacent venule, as well as recruitment of white blood cells (leukocytes).

or large reductions (Forsberg *et al.*) in one of these mediators, mouse mast-cell protease-6 (mMCP-6). Because the many different mast-cell proteases have distinct substrate specificities⁷, these findings have clear implications for potential mast-cell functions.

But regulating the enzyme content of cytoplasmic granules may not be all that heparin does in mast cells. Given its extraordinarily negative charge^{2,10}, heparin has long been thought to be a storage site in the mast-cell granule for the positively charged amines histamine and serotonin¹³. In accord with this view, Forsberg *et al.*⁵ report that the peritoneal mast cells in their NDST-2-deficient mice show a large — roughly 94% — reduction in cell-associated histamine. Humphries *et al.*⁴, however, report a more modest reduction of about 30%.

How can we explain this apparent discrepancy? One reason could be the ages of the mice studied. Mast cells can be long-lived, and, in both rats and mice, the volume of the cytoplasmic granules and histamine content of the peritoneal mast cells increase with age¹⁴. This may lead to progressively larger differences in the histamine (or mMCP-6) content of the NDST-2-deficient cells compared with normal cells. Another reason could be technical differences in the methods used to quantify mast cells. Finally, the two strains of NDST-2-deficient mice could have phenotypic differences owing to, for example, differences in the two targeting vectors, effects of the targeting on nearby genes, or the genetic backgrounds of the mice.

The functions of heparin stored in resting mast cells and that released into the tissues by activated mast cells are probably distinct. However, given that the consequences of mast-cell activation examined by Humphries *et al.*⁴ and Forsberg *et al.*⁵ — enhanced vascular permeability and recruitment of white blood cells (leukocytes) — were not substantially influenced by heparin, which mast-cell functions might be? First, activation of mast cells has been shown to lead to clotting in the adjacent connective tissue, where mast cells normally mature¹⁵ (Fig. 1b). Perhaps such clotting is enhanced in the NDST-2-deficient mice. Second, as well as being an anti-coagulant, heparin can bind certain cytokines, chemokines and growth factors, many of which are produced by mast cells^{6,7,16}. Finally, depending on the circumstances, heparin can either promote or suppress the proliferation and migration of cells involved in vascular development, atherosclerosis, wound healing and tissue remodelling². Heparin also regulates the distribution and enzymatic activity of certain mast-cell proteases^{7,17}.

Heparin may, then, have many functions beyond regulating the storage of other mediators in the mast-cell cytoplasmic granule. In other words, the production and initial characterization of NDST-2-deficient mice is not

the end of the heparin story, but the beginning of an exciting new chapter in our attempts to understand the biology of this fascinating, clinically useful — and still enigmatic — natural product. □

James L. Zehnder and Stephen J. Galli are at Stanford University School of Medicine, L235, 300 Pasteur Drive, Stanford, California 94305-5324, USA.
e-mail: sgalli@leland.stanford.edu

- Hirsch, J., Salzman, E. W., Marder, V. J. & Colman, R. D. in *Hemostasis and Thrombosis: Basic Principles and Clinical Practice* (eds Colman, R. D., Hirsch, J., Marder, V. J. & Salzman, E. W.) 1151–1163 (Lippincott, Philadelphia, 1994).
- Lane, D. A., Bjork, I. & Lindahl, U. (eds) *Heparin and Related Polysaccharides Adv. Exp. Med. Biol.* Vol. 313 (Plenum, New York, 1992).
- Rosenberg, R. D. & Bauer, K. in *Hemostasis and Thrombosis: Basic Principles and Clinical Practice* (eds Colman, R. D., Hirsch, J., Marder, V. J. & Salzman, E. W.) 837–860 (Lippincott, Philadelphia, 1994).
- Humphries, D. E. *et al. Nature* **400**, 769–772 (1999).
- Forsberg, E. *et al. Nature* **400**, 773–776 (1999).
- Galli, S. J. *N. Engl. J. Med.* **328**, 257–265 (1993).
- Huang, C., Sali, A. & Stevens, R. L. *J. Clin. Immunol.* **18**, 169–183 (1998).
- McLean, J. *Am. J. Physiol.* **41**, 250–257 (1916).
- Howell, W. H. & Holt, E. *Am. J. Physiol.* **47**, 328–341 (1918).
- Jorpes, J. E. *Circulation* **9**, 87–91 (1959).
- Lin, E. & Perrimon, N. *Nature* **400**, 281–284 (1999).
- Marcum, J. A., McKenney, J. B., Galli, S. J., Jackman, R. W. & Rosenberg, R. D. *Am. J. Physiol.* **250**, H879–H888 (1986).
- Uvnas, B. *Agents Actions Suppl.* **36**, 23–33 (1992).
- Hammel, L., Lagunoff, D. & Kruger, P. G. *Lab. Invest.* **59**, 544–554 (1988).
- Wershil, B. K., Mekori, Y. A., Murakami, T. & Galli, S. J. *J. Immunol.* **139**, 2605–2614 (1987).
- Boesiger, J. *et al. J. Exp. Med.* **188**, 1135–1145 (1998).
- Schwartz, L. B. *J. Allergy Clin. Immunol.* **86**, 594–598 (1990).

Condensed-matter physics

Real metals, 2D or not 2D?

Michelle Y. Simmons and Alex R. Hamilton

The distinction between metals and insulators appears simple — metals conduct electricity whereas insulators do not. Yet, for the past 25 years, arguments have raged over whether a two-dimensional (2D) system can be regarded as a real metal or not. In a 2D system — such as a very thin metal film, or the active region of many semiconductor transistors — the electrons are constrained to move in a plane of negligible thickness. Although these structures may conduct at room temperature, it was generally accepted that as the temperature (T) is reduced to absolute zero they would become

insulating. The celebrated discovery of the quantum Hall effect in 1980 demonstrated that it is possible to have metallic states in a 2D system that persist to $T = 0$ by applying a strong magnetic field¹. But it has never been clear what happens to these states when the magnetic field is turned off, and the existence of a 2D metal without a magnetic field (that is, $B = 0$) is still strongly debated. Recently, new evidence emerged indicating a transition from insulating to metallic behaviour at $B = 0$ in extremely pure 2D semiconductors. On page 735 of this issue², Hanein *et al.* report an experiment that suggests a link between this

Fluid dynamics

Lights, camera, drip

Scientists at the University of Chicago have been focusing a high-speed camera on dripping glycerine in the hope of catching the moment when the fluid drop breaks off. By taking 10,000 frames per second of a glycerine and water mixture dripping through a nozzle, they have captured images, such as the one shown here, of a drop at the point of snap-off.

According to an analysis of these images published in *Physical Review Letters* (83, 1147–1150; 1999), drops about to break free are self-similar — that is, the shape of the drop just before breaking looks the same at different times, if you rescale the axes. Understanding such fractal behaviour is important for the physics of mixing, because the quality of a dispersion — crucial when spray-painting cars, for example — depends on the way it breaks into drops.

Getting to grips with the mathematics

of this problem may be relevant to other areas of physics, such as theoretical studies into the gravity around black holes. When a fluid drop is about to break, a singularity develops, rather like the space-time singularity of a black hole. Simulating the infinite character of gravitational collapse remains a mathematical headache, and so studying a dripping tap may be an attractive alternative.

Sarah Tomlin



new metallic behaviour at $B=0$ and the quantum Hall metal, and raises the possibility that both share a common physical origin.

The notion that 2D metals cannot exist in the absence of a magnetic field dates back 20 years, when powerful 'scaling' theories indicated that any amount of disorder would trap electrons, so preventing conduction and the existence of a metallic state³. These arguments are based on the quantum-wave nature of the electrons, whereby a travelling electron wave can be scattered from impurities back to its starting point. If these returning waves interfere constructively, the electrons become localized in one place and are less able to diffuse through the solid. At high temperatures this effect is weak and the sample *appears* metallic. As the temperature is reduced quantum interference becomes more important, so that at absolute zero all the electrons are localized and completely unable to move. Low-temperature experiments with both thin-metal films⁴ and 2D sheets of electrons in field-effect transistors^{5,6} confirmed these theoretical predictions, and for nearly two decades it was generally accepted that there can be no 2D metal at $B=0$.

It is a historical coincidence that at about the same time as a consensus was being reached that no metallic states could exist in 2D systems, the quantum Hall effect was discovered. The classical Hall effect occurs when a current-carrying conductor is placed in a perpendicular magnetic field. This produces a Hall voltage across the conductor, which rises linearly as the magnetic field is increased. In contrast, the Hall voltage measured in 2D semiconductors at extremely low temperatures rises in steps with a series of abrupt transitions between well-defined plateaux at which the Hall voltage is precisely quantized. It is only possible to explain this quantization of the Hall voltage by the existence of both localized (insulating) and extended (metallic) electron states that persist to $T=0$. So, increasing the magnetic field reveals a series of insulator-metal transitions as the current-carrying electrons alternately find themselves in localized and extended states. Despite many years of intense study, it is still not clear what happens to these extended states as the magnetic field goes to zero. Some argue that, below some non-universal magnetic field, the extended states simply disappear; others suggest that, at low magnetic fields, they 'float' up to higher energies, becoming inaccessible to electrons at low temperatures⁷.

In 1994, strong evidence for a $B=0$ phase transition from an insulator to a metal was found in extremely low-disorder silicon field-effect transistors⁸, in apparent contradiction with the prevailing scaling theory. Similar behaviour was subsequently observed in other material systems, independent of the sign of the charge carriers, indicating that the metallic state is a universal

property of all low-disorder 2D systems. At present there is no theoretical consensus as to the nature of this unusual metallic phase, but experiments suggest that strong interactions between the charge carriers (not considered in the original scaling theory) and the spin of the electron (or positively charged hole) both play a role⁹.

Hanein and co-workers² have now carried out an experiment that relates the $B=0$ and quantum Hall metal-insulator transitions. By tuning the carrier density in a high-quality 2D GaAs hole system, they are able to alter the magnetic field at which the quantum Hall metal-insulator transitions occur. Their results are unique because they followed these transitions to much lower magnetic fields than in previous studies¹⁰. They find that the transition associated with the quantum Hall effect at high magnetic fields evolves continuously into the $B=0$ metal-insulator transition. This implies that the extended states in the quantum Hall regime do not simply disappear as the magnetic field is reduced, nor float up indefinitely in energy, but continue to $B=0$ with some finite energy.

These results are intriguing because they link the quantum Hall effect, which can be understood without considering electron-electron interactions, with $B=0$ metallic behaviour that is only found in strongly interacting systems. A number of questions remain. If the two metals share a common physical origin, it is difficult to reconcile the fact that the $B=0$ metal is destroyed by applying a magnetic field *parallel* to the 2D plane^{11,12}, whereas the extended states in the

quantum Hall regime are not. Furthermore, although some theories attribute the $B=0$ metal to a new, many-body ground state, certain experiments suggest that it may simply be a finite-temperature spin-related scattering effect¹³. If this is the case, then other effects, such as phase-coherent localization, may reappear at even lower temperatures. A great deal of work remains to be done before we can finally answer the simple question: is it possible for a 2D system to be a real metal at $B=0$? □

Michelle Y. Simmons and Alex R. Hamilton are at the Semiconductor Nanofabrication Facility and National Pulsed Magnet Laboratory, School of Physics, University of New South Wales, Sydney 2052, Australia.

e-mails: Michelle.Simmons@unsw.edu.au

Alex.Hamilton@unsw.edu.au

1. von Klitzing, K., Dorda, M. & Pepper, M. *Phys. Rev. Lett.* **45**, 494–497 (1980).
2. Hanein, Y. et al. *Nature* **400**, 735–737 (1999).
3. Abrahams, E., Anderson, P. W., Licciardello, D. C. & Ramakrishnan, T. V. *Phys. Rev. Lett.* **42**, 673–676 (1979).
4. Dolan, G. J. & Oshroff, D. D. *Phys. Rev. Lett.* **43**, 721–724 (1979).
5. Bishop, D. J., Tsui, D. C. & Dynes, R. C. *Phys. Rev. Lett.* **44**, 1153–1156 (1980).
6. Uren, M. J., Davies, R. A. & Pepper, M. *J. Phys. C* **13**, L985–L993 (1980).
7. Khmel'nitskii, D. E. *Phys. Lett.* **106A**, 182–183 (1984).
8. Kravchenko, S. V., Kravchenko, G. V., Furneaux, J. E., Pudalov, V. M. & D'Iorio, M. *Phys. Rev. B* **50**, 8039–8042 (1994).
9. Rice, M. *Nature* **389**, 916–917 (1992).
10. Kravchenko, S. V., Mason, W., Furneaux, J. E. & Pudalov, V. M. *Phys. Rev. Lett.* **75**, 910–913 (1995).
11. Simonian, D., Kravchenko, S. V., Sarachik, M. P. & Pudalov, V. M. *Phys. Rev. Lett.* **79**, 2304–2307 (1997).
12. Pudalov, V. M., Brunthaler, G., Prinz, A. & Bauer, G. *JETP Lett.* **65**, 932–937 (1997).
13. Murzin, S. S., Dorozhkin, S. I., Landwehr, G. & Gossard, A. C. *JETP Lett.* **67**, 113–119 (1998).

Marine biology

No hibernation for basking sharks

Daniel Weihs

The phrase 'gentle giants' and the word sharks do not usually appear in the same sentence — sharks normally evoke the image of a bloody, tooth-filled mouth as in the film *Jaws*. But the biggest species of selachians (sharks) are, in fact, placid, slowly moving grazers that feed by filtering plankton. They swim with a widely opened mouth, engulfing small prey and the water they are in, and using long, bristle-like projections on the gill arches to remove the prey.

The only species of shark that uses this method of feeding — known as ram filter feeding — exclusively is the basking shark (*Cetorhinus maximus*). Because these sharks have to swim with a wide gape (Fig. 1, overleaf) that increases drag and, hence, the energetic cost of feeding, a long-standing theory has been that they may migrate (and even hibernate¹) during winter, when the concentration of plankton falls below a

relatively high threshold². But in *Proceedings of the Royal Society*, David Sims³ combines behavioural observations and theoretical calculations to show that the threshold for gainful filter feeding is probably much lower for the basking shark, implying that hibernation and migration are not necessary.

Actually, what Sims shows is that we probably do not need complicated explanations such as hibernation. These complex models came about because original investigators assumed the energy content of copepods (the main constituent of the planktonic soup ingested by basking sharks) to be three orders of magnitude larger than currently accepted values. This large overestimation did not result in absurd conclusions (which, presumably, would have alerted previous investigators to the error) because initial estimates of the energy cost of swimming were also between one and two orders of magnitude too big.

The fact that hibernation is probably not a part of the basking sharks' annual cycle may solve an irksome question — where and how do they hibernate? The meagre literature on basking sharks contains no observations of hibernation. Given that they are large, pelagic fish, some reaching lengths of over 10 m, surely they would have been observed if they hibernated close to the continental shelf. If, on the other hand, they hibernated in mid-waters in the open ocean (assuming this is possible at all), the sharks would be exposed to attack by predators or scavengers. At the very least they would have been colonized by symbionts and parasites, which attach themselves to immobile objects. Attacks of either kind, although probably not fatal owing to the size of these sharks, would result in tell-tale signs on their bodies. No signs of such depredations have been reported.

But if they don't hibernate, where do most full-sized sharks go during the winter? The lower energetic requirements for swimming, calculated by Sims³ using up-to-date information and theory, help to justify the assumption that basking sharks migrate to deep waters. However, feeding in deep waters is complicated by competition from more agile cephalopods and fish, and the lower feeding threshold may allow the sharks to use feeding grounds not frequented by potential competitors.

An important lesson from this story is that classical 'results' need to be checked and rethought in the light of current knowledge. This is especially relevant in studies of marine biology, because the open sea has not been extensively studied — it has been said that mankind knows more about the topography of the Moon than about the bottom of the ocean. However, both the hydrodynamic theory of how fish swim and our understanding of the biochemistry behind the ingestion and use of food have improved over the past

three decades. We can be confident in Sims' results because there are now well-established theoretical calculations for swimming thrust and energy requirements⁴, and these have been shown to fit a large range of fish sizes⁵, including several large species of shark.

Much still needs to be done. First, we need to search for the sharks' deep-water winter habitats, and to study their behaviour in these habitats. This can be facilitated using archival tags, which record parameters such as temperature, depth and swimming speed. Although such tags are still under development, some models can detach from the fish at predetermined times and transmit the data to satellites overhead. Second, in arriving at his results, Sims had to make several assumptions that need to be verified. These include catch efficiencies of the gaping mouth, the shape and intake of the ram filter feeding system, assimilation costs and metabolic efficiencies.

Here there is an interesting analogy to what used to be known as Gray's paradox. In the 1930s, Sir James Gray of Cambridge, one of the leading experimental biologists of this

century, estimated the efficiency of dolphin muscle. He did this by analogy to human muscle, and concluded that the swimming speeds achieved by dolphins cannot be accounted for by standard hydrodynamics. This resulted in a decades-long search for 'the dolphin's secret'. Many important discoveries and developments resulted from this search, including advances in hydrodynamic theory, and the study of flexible skin and polymer drag reduction. Finally, in the 1970s, when accurate data were used, the (by then improved) hydrodynamic theory was found⁴ to be sufficient to describe cetacean swimming. Studying sharks may yield similar dividends. □

Daniel Weihs is at the Technion — Israel Institute of Technology, Haifa 32000, Israel.

e-mail: dweihs@technion.ac.il

1. Matthews, L. H. *New Scientist* **280**, 756–759 (1962).
2. Parker, H. W. & Boeseman, M. *Proc. Zool. Soc. Lond.* **124**, 185–194 (1954).
3. Sims, D. W. *Proc. R. Soc. Lond. B* **266**, 1437–1443 (1999).
4. Webb, P. W. & Weihs, D. *Fish Biomechanics* (Praeger, New York, 1983).
5. Weihs, D. in *Principles of Animal Design* (eds Weibel, E. R., Taylor, C. R. & Bolis, L.) 121–128 (Cambridge Univ. Press, 1998).

Neurobiology

Cognition by a mini brain

Randolf Menzel and Martin Giurfa

We often expect little from small brains and cognitive miracles from big ones. But all brains were once small, in both evolution and development, reaching their respective levels of cognitive function only gradually. Elementary forms of cognition might, then, be present in small brains, and these could tell us what is really gained when brains become bigger. For example, the fruitfly *Drosophila melanogaster* (the geneticist's delight and the neurophysiologist's torment) has a particularly small brain. Very little is expected of this tiny brain, and, indeed, few impressive behavioural capacities have been found so far — a range of innate response patterns, some artistic sexual foreplay, and rather limited forms of learning^{1,2}. But new discoveries bring exciting prospects, using both molecular genetics and behavioural studies to analyse the fruitfly's elementary cognitive functions.

On page 753 of this issue, Liu *et al.*³ report their results on visual learning by *Drosophila* and its underlying neuronal substrate. They show that individual flies can do quite complex tasks. The authors first conditioned flies to associate visual patterns (the 'conditioned' stimulus) with the presence or absence of heat (the 'unconditioned' stimulus). The idea is that the animals should fly towards the appropriate patterns to avoid dangerous levels of heat. Their behaviour is 'operant', because the flight course that they choose determines delivery of the heat. The

experiments were done under particular illumination conditions, which form part of the general 'context' in which the associations are established.

The authors next showed that the flies could 'generalize' this trained response to several other, different environmental contexts. As contexts are ill-defined stimuli, comprising individual features from many modalities, the change in context that the authors introduced was a change in the illumination between training and test: between white and monochromatic broad-band light; between two monochromatic broad-band lights; and between constant white light and white light interspersed with 'dark flashes' (light is switched off for 200 ms). These changes did not affect the performance of the flies. These results show that context generalization — rather than context specificity — guides the insect's learning. But when the authors impaired the fly's normal brain function by eliminating the mushroom bodies (a central brain structure), they found that retention of the trained pattern was strictly bound to the context during learning, and that the flies did not generalize to other contexts.

The study by Liu *et al.*³ sheds light on two issues: the function of a central brain structure on an elementary form of cognition, and the role of context in learning. Mushroom bodies are prominent paired structures in the insect brain (Fig. 1), and are considered



Figure 1 Feeding time. The basking shark (*Cetorhinus maximus*) feeds on plankton, which it filters from sea water using bristle-like projections on the arches of its gills. In winter the concentration of plankton is reduced below a threshold level, and sharks were thought to escape starvation by hibernating or migrating. But a study by Sims³ indicates that the threshold of plankton needed by the sharks is much lower than previously estimated, and that the sharks do not need to hibernate or migrate in winter.

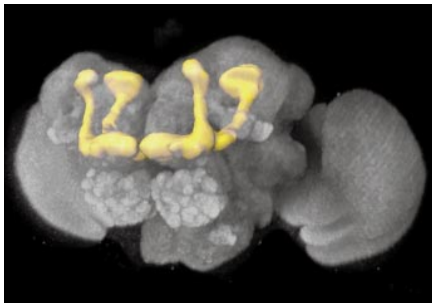


Figure 1 Mushroom bodies from the fruitfly *Drosophila melanogaster*. These paired structures are thought to be an information-processing centre, and Liu *et al.*³ now show that they are involved in context generalization during learning.

to be a high-order information-processing centre^{4,5}. In *Drosophila* they are related to olfactory processing and formation of olfactory memories¹, but they do not seem to be involved in visual processing and memory⁵.

The authors used three different techniques to remove the mushroom bodies (an *mbm* mutant lacking mushroom bodies; selective pharmacological ablation using the cell poison hydroxyurea during the first larval stage; and reverse-genetic ablation by enhancer-driven tetanus toxin), and showed that visual input affects the mushroom bodies. They suggest that the mushroom bodies do a dynamic analysis of the structure of the environmental stimuli, such that their temporal trace is compared to that of the unconditioned stimulus. (A similar function has been proposed for the honey-bee mushroom body, based on rather different data⁴.) Only the visual stimuli that change in a synchronous way with the heat stimulus can be useful predictors of dangerous heat. Other stimuli have a temporary structure, different from that of the unconditioned stimulus, so they are classified as 'context'. Flies that lack mushroom bodies cannot generalize their learned response to a different context because they are unable to dissociate the visual patterns from the illumination under which they were trained. In these flies, a compound between context and conditioned stimulus is formed. So, when the context changes, the compound changes too, impeding generalization to a new context.

The mushroom body seems to have three general functions. First, it separates stimuli according to their temporal structure. Second, it evaluates these stimuli with respect to the unconditioned stimulus (context or predictor). Finally, during the retention phase, it grades the stimuli according to predictive power. Such a proposal for a structure–function relationship is similar to a hypothesis by Solomon and Moore⁶ for the hippocampus in mammals — poorly associated conditioned stimuli are 'tuned out', and a hierarchy of stimulus predictiveness is reached. In this sense, the mushroom body may allow

selective attention, an attribute not yet experimentally accessible in insects.

Liu and colleagues' study also provides a fresh view of context dependence, an old issue in learning psychology. The textbooks tell us that animals can learn different meanings for the same stimulus when it is presented in different contexts. They maintain that contextual and conditioned cues appear to an animal as separate sensory entities, which may be associated in different ways⁷. *Drosophila*, on the other hand, seem to use a different system in which contextual cues and conditioned stimuli are filtered by the mushroom bodies. If this idea is correct, *Drosophila* may never have the flexibility to establish different, conditional relationships between contextual and conditioned stimuli.

But does the *Drosophila* mushroom body indeed act only on a basic sensory-processing level, as suggested by Liu *et al.*? Flies that lack the mushroom body perform tasks according to strict training conditions, indicating that sensory processing and association are done and stored outside the mushroom body, probably in the visual ganglia. So, the mushroom bodies might be related to the organization of multiple memories. It is possible that storage or retrieval initiates a search routine for additional conditional forms of memory traces — such as a trace indicating one stimulus as the contextual cue and the other as the conditioned stimulus. Such a trace would reside in the mushroom bodies, and would carry the information about whether the context or the conditioned stimuli were stable or variable. Flies deficient in the mushroom bodies may lack this trace, so would behave exactly as Liu *et al.* describe.

Although Liu and colleagues' data do not require such an elaborate model, their experiments have not yet tested it. The next steps may be to try and uncouple the function of mushroom bodies during acquisition, storage and retrieval, and to test combinations of contextual and conditioned cues in a balanced fashion. *Drosophila* will be an ideal organism in which to do these experiments, given the tools available for inducible knock-out mutants. In any case, the mini brain of this little fly is now well qualified as a model for basic cognitive functions. □

Randolf Menzel and Martin Giurfa are at the Freie Universität Berlin, Königin-Luise-Strasse 28/30, D-14195 Berlin-Dahlem, Germany.
e-mails: menzel@neurobiologie.fu-berlin.de
giurfa@neurobiologie.fu-berlin.de

1. Dubnau, J. & Tully, T. *Annu. Rev. Neurosci.* **21**, 407–444 (1998).
2. Greenspan, R. J. *Neuron* **15**, 747–750 (1995).
3. Liu, L., Wolf, R., Ernst, R. & Heisenberg, M. *Nature* **400**, 753–756 (1999).
4. Erber, J., Homberg, U. & Gronenberg, W. in *Arthropod Brain: Its Evolution, Development, Structure, and Functions* (ed. Gupta, A. P.) 485–511 (Wiley, New York, 1987).
5. Heisenberg, M. *Learning Mem.* **5**, 1–10 (1998).
6. Solomon, P. R. & Moore, J. W. *J. Comp. Physiol. Psychol.* **89**, 1192–1203 (1975).
7. Holland, P. C. & Bouton, M. E. *Curr. Opin. Neurobiol.* **9**, 195–202 (1999).

Daedalus

Superficial complexity

Metallic mercury has a high surface tension. But make it an electrode in an ionic solution, apply a voltage, and ions crowd its surface in surfactant fashion. Its surface tension drops dramatically. Non-ionic additives can reduce the tension even further. Daedalus is now taking this electro-capillary effect to extremes, using other metals.

He has set up a heated electrochemical cell with a molten metal electrode at the bottom, and a layer of molten salt above it as electrolyte. The vastly greater ion density of the melt will reduce the surface tension of the metal interface far more. Indeed, with a high enough voltage, a well-chosen salt, and some cunningly selected non-ionic additives, the surface tension of the metal could be driven negative.

A liquid with negative surface tension should increase its surface area vigorously. In the simplest case it might break into separate droplets; but not in Daedalus's system. Once detached from the bulk, a globule would cease to share its voltage. It would revert to positive surface tension, and reattach itself. Instead, the metal surface will fold into some nightmarishly complex fractal, dense with tiny corrugated corrugations of vast total surface area, but still all one surface.

Powerful applications beckon for this new metal treatment. Catalysis, heat exchange, electrode forming, battery and capacitor manufacture, all cry out for metals of the highest possible surface area. Daedalus is planning a 'continuous fractalization bath', in which a strip of base material, coated with the metal to be fractalized, runs at high speed through his molten-salt cell. The coating melts, acquires negative surface tension, blooms into vast surface area, and sets solid again in a cooler region of the bath downstream. Its adhering salt coating is washed off with water, and it is reeled up.

The pilot studies will use platinum, whose catalytic surface is crucial to so much chemistry. A possible salt for the job is potassium hydrogen sulphate; its melt cleans platinum so splendidly that it may be a natural surfactant for the metal. But Daedalus's fractal metal surfaces should find many other uses. In particular, they will be ideal for gluing and bonding. Unlike surfaces made by grinding or scoring, they will have re-entrant cavities and protrusions on all scales, and will keep an applied paint or resin immovably in place. Even PTFE-coated saucepans will at last retain their slippery coating.

David Jones

Obituary

Rolf Landauer (1927–99)

Head and heart of the physics of information

It is well known that scientific progress requires intellect. Less advertised is science's need for conscience. Intellect creates new ideas; conscience examines them for consistency and correctness. In other words, a scientific field requires not only 'head' but 'heart' as well. Rolf Landauer was both head and heart to the field of physics of information. He died of cancer on 27 April this year at the age of 72.

Landauer based his research on a simple rule: information is physical. That is, information is registered by physical systems such as strands of DNA, neurons and transistors; in turn, the ways in which systems such as cells, brains and computers can process information is governed by the laws of physics. Landauer's work showed that the apparently simple and unproblematic statement of the physical nature of information had profound consequences.

Born in Stuttgart, Landauer was raised in New York from the age of eleven. He received both undergraduate and graduate degrees from Harvard, where he worked under Léon Brillouin and Wendell Furry. In 1952 he joined IBM, where he was employed for the rest of his life. In the late 1950s, after constructing an influential theory of electrical conductance based on the scattering of electrons, Landauer turned his attention to the physics of computation.

In 1961, Landauer discovered that logical operations that get rid of information, such as erasure, necessarily require the dissipation of energy. Erasure transforms information from an accessible to an inaccessible form, known as entropy, whereas logical operations that can be reversed do not lead to a rise in entropy. Landauer also identified the minimum amount of entropy increase required for each bit erased as $k \ln 2$, where k is Boltzmann's constant. Landauer's principle transformed the physics of information by identifying the basic trade-off between information and physical quantities.

In the 1980s, Landauer's colleague at IBM, Charles Bennett, applied his principle to provide a self-consistent solution to the century-old problem of Maxwell's demon. In his research into the molecular nature of heat in the mid-nineteenth century, James Clerk Maxwell had noted that a hypothetical being that



could acquire information about the velocities of individual molecules of a gas could shepherd fast molecules in one direction and slow molecules in another, thereby decreasing the gas's entropy. But the second law of thermodynamics says that entropy does not decrease. Because of its pernicious effect, William Thompson and Lord Kelvin dubbed this being 'Maxwell's demon'. In order to preserve the second law of thermodynamics, scientists earlier this century postulated that getting and processing information must necessarily involve an increase in entropy. Landauer's principle shows that, whereas Maxwell's demon can, in principle, use information about the gas to do work without increasing entropy, when it erases that information it must pay a price of $k \ln 2$ in entropy per bit, exactly cancelling the advantage it obtained in the first place.

In the 1970s, working from Landauer's result, Bennett (and independently Ed Fredkin and Tom Toffoli of MIT) showed that all computation could, in principle, be carried out in a logically reversible fashion: as a result, computation does not require dissipation. The fact that modern computers dissipate large amounts of heat arises from practical engineering concerns rather than from physical necessity. The possibility of reversible computation was central to Paul Benioff's discovery at Argonne National laboratory in 1980 that quantum systems could perform computations in a coherent fashion, a result that led to quantum computation.

Although he may be considered the godfather of quantum computation, Landauer was fundamentally critical of the field. During the 1960s, he was a senior manager in IBM's research division, and played a role in many of the laboratory's successes, including the development of the injection laser and of large-scale integrated memory circuits. During his

career, he witnessed the failure of many promising technologies, such as early attempts to construct computers using Josephson junctions and tunnel diodes. Having seen them come and go, Landauer was sceptical of new computer technologies, noting that most proposals gravely underestimated the difficulty in constructing a viable device. Over the past 30 years, he wrote a series of articles pointing out deficiencies in various alternative computing proposals, including quantum computation. When a researcher failed to pay adequate attention to his published admonitions, he would warn them in person. For example, he suggested that all papers on quantum computing should carry a footnote: "This proposal, like all proposals for quantum computation, relies on speculative technology, does not in its current form take into account all possible sources of noise, unreliability and manufacturing error, and probably will not work."

Landauer's footnote is, of course, correct. In any field of technology, let alone one that attempts to store information on subatomic particles, most ideas will probably not work. It is by identifying these wrong ideas that scientists and engineers winnow out those few ideas that are potentially right. Yet, despite the importance of mistakes in science, most published papers report work that is potentially right, rather than provably wrong. In such an environment, it is important for a field to have a mechanism for remembering past mistakes and for preventing them from recurring. In the field of physics of information, Landauer supplied such a mechanism by functioning as the field's conscience.

Landauer's bracing criticism had a highly positive and healthy effect on quantum computing. His critiques were always accompanied by advice, support and suggestions for making the suspect proposal work. Researchers explicitly addressed his concerns and developed methods for coping with noise, unreliability and manufacturing error, to the extent that, by the end of his life, although he still believed that large-scale quantum computers would not be built, he admitted that they were considerably more plausible than they had been a few years before. It is not yet clear what the future physical basis for computation will be, but, whatever its successes, they will be in no small part due to Rolf Landauer.

Seth Lloyd

Seth Lloyd is in the Department of Mechanical Engineering 3-160, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA. e-mail: slloyd@mit.edu

Rules of the game of doing science

Scientific good practice – disinterested, communal, universal – is not ‘blobby’ idealism, but a social framework to which researchers must conform if they are to prosper. Now, this contract with society is up for renegotiation.

John Ziman

Nineteen-forty-two was a pretty awful year. Democracy had its back to the wall. Science was a sinew of war. But this was also the year when it was shown that science and democracy are indissolubly united.

Robert Merton was a whiz-kid professor of sociology at Columbia University, New York. He was an expert on the place of science in society. So he gathered some of his ideas into a little paper on “Science and Society in a Democratic Order”. Its appearance, in the first issue of an obscure journal, is the event I celebrate (*J. Legal and Political Sociology* 1, 115–262; 1942). Not that I — nor, I suspect, many others — read this paper at the time. It didn’t come my way until nearly 30 years later, perhaps as an exchanged offprint, perhaps in a collection of the writings of this now famous scholar. By then its title had become “The Normative Structure of Science”. As sometimes happens, a conceptual skeleton constructed to stiffen a slightly floppy argument had proved more enduring than its empirical flesh. Merton’s X-ray eye had detected the internal social framework that gives science its strength.

In the then sociological fashion, he described science as an “ethos”, held together by “norms”. Research results are “communal” — they belong to the whole scientific community. They must be “universal” — independent of class or creed. They must be presented “disinterestedly”, and be subject to “organized skepticism”. At first, these sound like blobby ideals, mainly suitable for public panegyrics. But Merton saw them functionally, as the regulatory principles of a way of life. These are the rules of the game of doing science, which every player is forced to obey if she or he is to stay on the field.

The Mertonian norms are not just unattainable personal attitudes. They are embodied in innumerable social conventions and mundane practices. These define and constrain our conduct as scientists. Publish — or perish. Face up to the demands of peer review. Cite generously and meticulously. Reward originality and priority of discovery. Present your work impersonally. Exclude *ad hominem* jibes. And so on.

These requirements are often irksome. We may sometimes be tempted to keep profitable ideas secret, skimp on measurements, fudge data, pirate the work of others, ignore criticism, boost ourselves or vilify our opponents. But our good name is at stake. Without personal credibility, our research would

be disregarded, and jobs would no longer come our way. So we strive to perform as expected, and as we expect of our colleagues. The logic of life forces us to think and act scientifically, until it becomes second nature.

Merton’s novel insight was that science — he meant ‘pure’ science, for industrial research and technology work differently — is not just the activity of a community of like-minded individuals. It is a distinctive institution, with a distinctive culture. Its everyday practices dovetail into a compelling social framework. This framework supports a ‘method’, an attitude of mind, a profession, a body of knowledge. He made us realize that science is driven, shaped and honed as much by its internal sociology as by its philosophy or psychology.

In 50 years, the sociology of science has grown into an established academic discipline. Unfortunately, the focus of this discipline has shifted. Sociology seems almost against science as it strives to put it in its place. Well, of course, scientific knowledge is not uniquely true, science is not the only pebble on the social beach, and scientists, like traffic wardens and airline pilots, are ordinary folk too. These are realities that we should all humbly accept.

Yet science remains a peculiar institution, with its own way of doing things. Love it or loathe it, we need to understand just what

makes it tick. Is extreme specialization, for example, an essential cog in the social machine? What winds the clockwork — social imperatives, coordinated curiosity or competitive ambition? How can anybody tell whether it is working properly? Whose hands are on the lever for adjusting its output?

What is more, science is no longer what it was when Merton first wrote about it. The bureaucratic engine of policy is shattering the traditional normative frame. Big science has become a novel way of life, with its own conventions and practices. What price now those noble norms? Tied without tenure into a system of projects and proposals, budgets and assessments, how open, how disinterested, how self-critical, how riskily original can one afford to be?

There is no going back to that world we have lost. Anyway, science has never had it so good as in this past half-century, and is still going great guns. But soon it will be offered a new contract with society. To renegotiate that contract with its eyes open, on even terms, science will need to understand itself much better. That understanding is going to require, not adherence to an obsolete ethos, but a sharp but sympathetic sociological self-analysis. That is the unfinished business that Merton’s little paper began. □

John Ziman is at 27 Little London Green, Oakley, Aylesbury, Buckinghamshire HP18 9QL, UK.



“Scientists have extended the life of the fruit fly.”

THE NEW YORKER/GEORGE BOOTH/CARTOONBANK.COM

Chemokine control of HIV-1 infection

Chemokines are proinflammatory cytokines that attract and activate specific types of leukocyte¹. There are two main chemokine families, based on the position of the first two cysteine residues: the CC and the CXC chemokines¹. Chemokines mediate their effects through interactions with seven-transmembrane-spanning glycoprotein receptors coupled to a G-protein signalling pathway¹. Chemokine receptors normally undergo a ligand-mediated homodimerization process, which is required for Ca^{2+} flux and chemotaxis². Here we show that in the chemokine response it is possible for heterodimerization, rather than homodimerization, to occur between a mutant form of the CCR2 receptor (the CCR2V64I receptor), which helps to delay the development of AIDS in HIV-1-infected individuals, and the CCR5 or CXCR4 chemokine receptor, which are used by HIV to gain entry into cells. These results may explain why AIDS takes longer to develop in HIV-1-infected individuals carrying the CCR2V64I mutation³.

The discovery that certain chemokine receptors can also act as receptors for HIV-1, HIV-2 and simian immunodeficiency virus has shed light on the mechanisms underlying viral entry and tropism⁴. The virus strains responsible for transmission, which are the predominant virus type isolated from asymptomatic HIV-positive individuals, use CCR5 as a receptor, whereas viruses that emerge later in the course of infection use CXCR4 to gain entry into the cell, either instead of or in addition to CCR5 (ref. 5).

The CCR2V64I polymorphism of CCR2, in which a valine at position 64 is replaced by an isoleucine, occurs at an allele frequency of 10–25%, depending on the ethnic group of the population, and is associated with a delay of 2–4 years in the progression to AIDS³. Relatively few viral strains have been reported that can use CCR2 in conjunction with CD4 to infect cells, so the mechanism underlying this protective effect is not clear⁶.

After binding to their specific receptors, chemokines induce receptor homodimerization and activate the receptor-associated JAK kinase, possibly by transphosphorylation of tyrosine residues^{2,7}. This creates SH2 docking sites, leading to the recruitment of STAT transcription factors. In another member of the seven-transmembrane-spanning receptor family, the response to the neurotransmitter GABA requires heterodimerization of the GBR1 and GBR2 receptors⁸. The interaction between GBR1 and GBR2 seems to be essential for the activation of potassium channels.

One possibility is that the protective effect of CCR2V64I is due to its ability to

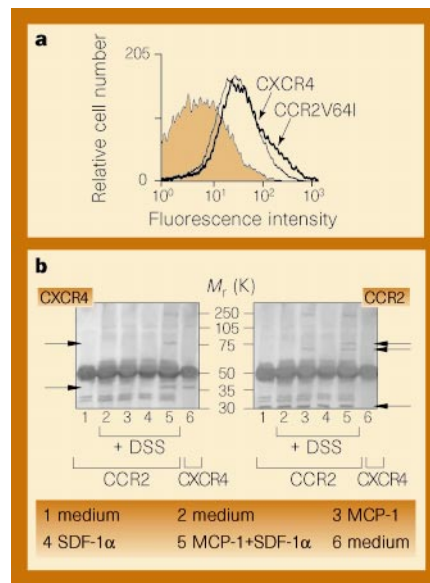


Figure 1 Ability of the mutant CCR2V64I receptor to form heterodimers with the CXCR4 receptor. **a**, HEK-293 cells, which constitutively express CXCR4, were transfected with the expression vector for CCR2V64I and stained¹⁰ with a monoclonal antibody specific for CCR2 and CXCR4, using isotype-matched antibody as a control. **b**, CCR2V64I-transfected HEK-293 cells were stimulated with chemokines (10 nM, 5 min at 37 °C) and, where indicated, crosslinked with 1 mM DSS. Cell lysates were immunoprecipitated with anti-CCR2 antibody, electrophoresed and transferred to nitrocellulose membranes. The western blot was analysed with anti-CXCR4 antibody; as a positive control, lysates from the same cells were immunoprecipitated with anti-CXCR4 antibody (left). The membrane was stripped and reblotted with anti-CCR2 antibody as a protein loading control (right). Arrows indicate the monomer and dimer.

heterodimerize with the CCR5 and/or CXCR4 receptor. We tested this idea by transfecting HEK-293 cells, which constitutively express the CXCR4 receptor, with the CCR2V64I mutant (Fig. 1a) and monitoring the response to the chemokines MCP-1 and SDF-1 α . Equivalent expression of both receptors was verified by flow cytometry analysis and by their ability to respond in chemotaxis and in Ca^{2+} flux to either MCP-1 or SDF-1 α . In these cells, MCP-1 and SDF-1 α sensitize responses to the homologous, but not the heterologous, chemokine (results not shown).

To test for heterodimerization of the CCR2V64I and CXCR4 receptors following MCP-1 and SDF-1 α binding, cells were first crosslinked by using disuccinimidyl suberate and then lysed. Anti-CCR2 immunoprecipitates of cell lysates were developed in a western blot by using anti-CXCR4 or anti-CCR2 antibodies. In CCR2-derived immunoprecipitates, CXCR4 receptors

were observed after MCP-1 and SDF-1 α activation, but not in the presence of either chemokine alone (Fig. 1b, left). The membrane was then stripped and reprobed with anti-CCR2 antibody. As expected, MCP-1 was found to induce dimerization of the CCR2 receptor (Fig. 1b, right), as shown by the presence of dimers or complexes of higher relative molecular mass.

In contrast, CCR2/CXCR4 co-transfected receptors do not undergo heterodimerization in response to stimulation by MCP-1 and SDF-1 α (not shown). This indicates that the isoleucine at position 64 is required to promote heterodimerization between the CCR2V64I and CXCR4 receptors. Isoleucine is also found at position 64 in the CCR5 receptor. In similar experiments with CCR5-transfected HEK-293 cells, we also detected CCR2V64I in CCR5 immunoprecipitates when cells were stimulated simultaneously with the chemokines RANTES and MCP-1 (data not shown). Finally, the CCR2V64I mutant receptor and the wild-type CCR2 receptor also heterodimerize with the CCR5 receptor (results not shown).

Our data showing that CXCR4 can dimerize with the CCR2V64I mutant, but not with wild-type CCR2, indicate that Val 64 may be critical for dimer stabilization. We have previously suggested that chemokine-receptor dimerization may help to prevent HIV-1 infection⁹. The ability of CCR2V64I to dimerize with CXCR4 may reduce the amount of CXCR4 on peripheral blood mononuclear cells. Together with the ability of CCR2V64I to heterodimerize with CCR5, this might explain the delayed progression to seroconversion in HIV-infected donors carrying this mutation¹.

These findings may help us to understand the molecular mechanisms responsible for chemokine-receptor signalling and to explain why some chemokines block HIV-1 infection. One implication of receptor clustering following chemotactic responses is that the activity of one receptor could influence that of its neighbours, so a ligand might be able to act *in trans* on a receptor for which it is not specific.

Mario Mellado, José Miguel Rodríguez-Frade, Antonio J. Vila-Coro, Ana Martín de Ana, Carlos Martínez-A.

Department of Immunology and Oncology,
Centro Nacional de Biotecnología/CSIC,
Universidad Autónoma de Madrid,
Campus de Cantoblanco,
28049 Madrid, Spain
e-mail: cmartineza@cnb.uam.es

1. Baggiolini, M. *Nature* **392**, 565–568 (1998).

2. Rodríguez-Frade, J. M. et al. *Proc. Natl. Acad. Sci. USA* **96**, 3628–3633 (1999).

3. Smith, M. W. *et al.* *Science* **277**, 959–965 (1997).
4. Berger, E. A., Murphy, P. M. & Farber, J. M. *Annu. Rev. Immunol.* **17**, 657–700 (1999).
5. Cairns, J. S. & D'Souza, M. P. *Nature Med.* **4**, 563–568 (1998).
6. Rucker, J. *et al.* *Cell* **87**, 437–446 (1996).
7. White, J. H. *et al.* *Nature* **396**, 679–682 (1998).
8. Vila-Coro, A. J. *et al.* *FASEB J.* (in the press).
9. Rodríguez-Frade, J. M. *et al.* *J. Cell Biol.* **144**, 755–765 (1999).
10. Rodríguez-Frade, J. M. *et al.* *J. Clin. Invest.* **100**, 497 (1997).

Climate variability and crop yields in Europe

Hulme *et al.*¹ make a valid distinction between the relative importance for European water runoff and wheat yields of human-induced climate change and naturally occurring climate variability, when considered over timescales of several decades. They use the HadCM2 climate model to show that, in many European cases, the human-induced climate change 'signal' could be swamped by naturally occurring climatic 'noise'.

However, the effect of any climate variability signal also depends on the temporal scale (for example, annual versus decadal) of such variability. For wheat crop yields, the relevant timescale for climatic variability is over years or even days, rather than decades, simply because wheat is an annual crop. Hulme *et al.* decided not to alter either inter-annual or inter-daily climatic variability.

By using an earlier global climate model (UKTR; ref. 2) than Hulme *et al.*, we have analysed the response of a wheat simulation model³ to scenarios of climate change with and without changes in the inter-annual variability of precipitation (intensity and occurrence) and temperature. We introduced changes in climate variability derived from UKTR using a stochastic weather generator⁴. To measure the effect of inter-annual climate variability, we studied mean grain yield and its coefficient of variation (CV) from 30 individual years of simulation (the same total period as used by Hulme *et al.*).

We looked at two of the same sites as Hulme *et al.* For Spain, the simulated mean for the baseline climate wheat yield was 5.6 tonnes per hectare, with a CV of 0.24 (Table 1). In the climate scenario without a change in inter-annual variability, mean yield fell slightly to 5.2 t ha⁻¹, with a CV of 0.23. With changes in climate variability, simulated mean yield dropped to 3.9 t ha⁻¹ but its CV more than doubled to 0.48, largely

because there were more prolonged dry spells over the vegetation period. The probability of producing yields of less than 3.5 t ha⁻¹ in the 'with variability' scenario was nearly 0.50, compared with about 0.10 for the baseline and 'without variability' scenarios. Such changes in annual yield variability would make wheat a risky crop to grow in Spain and have important economic and social consequences.

For the United Kingdom, changes in climate variability had little effect on either mean grain yield or its CV (Table 1). In contrast, Hulme *et al.* found no change in the range (and hence the variability) of simulated yields for these sites as a result of either multi-decadal natural or anthropogenic variability. The timescale of the imposed variability can therefore alter qualitatively whether or not variability has an effect.

We conclude that it is important when assessing the impact of climate change to differentiate between natural climate variability and anthropogenic climate change, as highlighted by Hulme *et al.*, but also to apply changes in climate variability at appropriate timescales.

John R. Porter*, **Mikhail A. Semenov†**

*Department of Agricultural Sciences,
The Royal Veterinary and Agricultural University,
2630 Taastrup, Denmark
e-mail: jrp@kvl.dk

†ACR Long Ashton Research Station,
University of Bristol, Bristol BS41 9AF, UK

1. Hulme, M. *et al.* *Nature* **397**, 688–691 (1999).
2. Murphy, J. M. & Mitchell, J. F. B. *J. Clim.* **8**, 57–80 (1995).
3. Jamieson, P. D. *et al.* *Eur. J. Agron.* **8**, 161–180 (1998).
4. Semenov, M. A. & Barrow, E. M. *Clim. Change* **35**, 397–414 (1997).

Hulme *et al.* reply — Porter and Semenov point out the importance of changes in climate variability on different timescales when assessing the significance of climate-change impacts. They are correct to assert that changes in inter-daily, inter-annual and multi-decadal climate variability are important, but our study¹ was intended to draw attention to natural multi-decadal climate variability and to indicate how it might obscure the identification of significant effects of anthropogenic climate change. It was therefore primarily a study of signal-to-noise ratios in impacts indicators, something that has not previously been attempted using low-frequency climate variability. The effects on crop yields of changes in inter-annual and inter-daily climate variability have been considered previously^{2–4}.

We do not agree with Porter and Semenov that the primary climate-variability timescale of relevance for wheat yields is the inter-annual. Changes in the daily variability of climate and in the frequency of extreme events can be very important for crop performance⁵, and daily climate variability can change independently of annual or decadal timescales. Porter and Semenov did not

perturb daily climate variability in their study². We think that decadal climate variability is also important for wheat yields: such low-frequency changes in climate (whether or not they are accompanied by changes in higher climate frequencies) can profoundly affect crop suitability and economic viability⁶ and alter the competitiveness of different agricultural products on these timescales.

The effect of anthropogenic changes in inter-annual and/or inter-daily climate variability on an impact indicator will depend on their magnitude with respect to natural changes (for example, for Europe, in relation to the North Atlantic Oscillation). This is also a signal-to-noise problem that can be explored using our methodology, but was not addressed by Porter and Semenov.

The challenge for those studying the impacts on climate is to examine the effects of changes in the full spectrum of climate variability, clearly distinguishing between anthropogenic and natural changes, and interpreting these effects in the light of other forces of environmental change. If these can be quantified more comprehensively than in the past, then we can provide decision-makers and resource managers with information that can better assist them to manage future risks arising from climatic change.

Mike Hulme*, **Paula Harrison†**,
Nigel Arnell‡

*Climatic Research Unit,
School of Environmental Sciences,
University of East Anglia, Norwich NR4 7TJ, UK
†Environmental Change Unit,
University of Oxford, Oxford OX1 3TB, UK

‡Department of Geography,
University of Southampton,
Southampton SO17 1BJ, UK

1. Hulme, M. *et al.* *Nature* **397**, 688–691 (1999).
2. Semenov, M. A. & Porter, J. R. *Agric. Forest. Meteorol.* **73**, 265–283 (1995).
3. Riha, S. J., Wilks, D. S. & Simeons, P. *Clim. Change* **32**, 293–311 (1996).
4. Mearns, L. O., Rosenzweig, C. & Goldberg, R. *Clim. Change* **35**, 367–396 (1997).
5. Harrison, P. A., Butterfield, R. E. & Downing, T. E. *ECU Res. Rep.* **9** (ECU, Oxford, 1995).
6. Parry, M. L. & Carter, T. R. *Clim. Change* **7**, 95–110 (1985).

A movement-sensitive area in auditory cortex

It is important to recognize sound patterns, regardless of their position and motion. The ability to locate sound sources and track their motion involves various levels of the auditory pathways^{1,2}. Motion and pattern analysis may first be spatially separated in the auditory cortex. We have examined this by using functional magnetic resonance imaging (fMRI) and find a higher-order field in the right auditory cortex that is

Table 1 Mean grain yield with and without changes in inter-annual climate variability

	Mean grain yield (t ha ⁻¹) (CV)	
	Spain	UK
Baseline climate	5.6 (0.24)	9.8 (0.06)
UKTR without changes in inter-annual variability	5.2 (0.23)	11.5 (0.08)
UKTR with changes in inter-annual variability	3.9 (0.48)	11.4 (0.09)

Simulated over a period of 30 years.

activated by sound motion significantly more than other fields of the auditory cortex. This area distinguishes whether a sound pattern is moving or stationary.

We compared activation of the auditory cortex by two acoustic stimuli with identical spectral-temporal patterns but with different binaural timing. The pattern consisted of a tonal carrier that was amplitude modulated (AM, 80% modulation depth, sine² waveform) at slow rates (0.08 to 0.2 Hz). In the 'movement' version, the envelope was 90° out of phase at the two ears, generating perception of a sound moving at varying speed in azimuth. In the 'stationary' version, the envelope was in phase at the two ears, generating perception of varying loudness of the same, but stationary, sound. The AM carrier, a frequency-modulated (FM) tone (1 to 6 kHz, 4 Hz sawtooth modulation), was designed to activate a broad range of frequency channels and was previously found to delineate various fields in human auditory cortex. Four such fields — T1a, T1b (AI), T2 and T3 — were distinguished by a combination of structural landmark delineation and differential functional activation³.

In ten right-handed subjects, three contiguous fMRI slices (4 mm thick) were obtained encompassing and parallel to the superior plane of the temporal lobe in both hemispheres. Electrodynamical headphones³ were used for stimulus presentation (root-mean-square 72 dB sound pressure level, SPL). For each slice, 108 functional images were collected by using a low-noise imaging sequence (48 dB SPL)³ at 3 tesla. The experiment consisted of six cycles comprising stationary FM, moving FM and silent phases. Attention was controlled by target detection (pauses of the AM lasting 2–4 seconds, right key press). Functional activation in individual subjects was assessed by correlations to the stimulus input functions. Activation differences across subjects were analysed by using intensity-weighted volumes of significantly activated volumes⁴.

We found activation by both the moving and stationary patterns with respect to silence in all four fields of auditory cortex in both hemispheres (Fig. 1a). Across-subject analysis contrasting the activation from the moving and stationary patterns revealed a large residual activity in only T3 on the right planum temporale (Fig. 1b,d). This residual activity was significantly different from that of the left T3 (Wilcoxon, $P=0.037$) and all other fields bilaterally (Fig. 1c).

Our results indicate that there is an area in the auditory cortex whose activity distinguishes whether a sound pattern is moving or stationary. Both the moving and stationary stimuli strongly activated various fields of the auditory cortex bilaterally in which this motion dependence

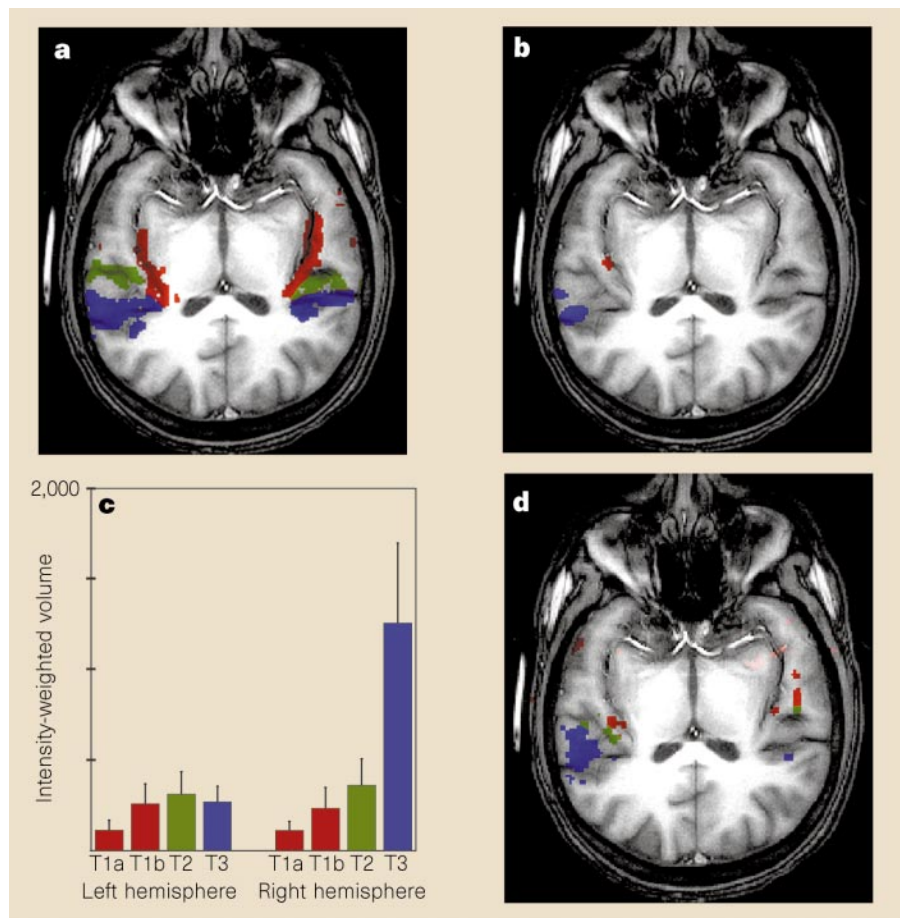


Figure 1 Nonspecific and specific activation by motion stimulus. **a**, Motion-induced activation compared with silence in a single subject mapped on the dorsal surface of the temporal lobe ($P<0.05$). Voxels in different fields are colour-coded (T1b(AI) and some T1a voxels more rostrally, red; T2, green; T3, blue). **b**, Correlation contrast activity difference from moving compared with stationary sound in the same subject ($P<0.05$). **c**, Residual intensity-weighted volume (volume in microlitres × percentage intensity) across subjects for moving compared with stationary stimuli in all fields. **d**, As **b**, but showing the average from all ten subjects mapped onto an individual brain.

was not seen. These data do not imply that T3 in the right hemisphere contains neurons that are movement selective (in the sense of responding only to movement or preferring a particular movement direction). But even if T3 were only a space map of position-sensitive neurons, its overall activation by motion is clearly higher than that of the other auditory cortical areas studied.

It is interesting that the significant effect was found only on the right planum temporale, because an area sensitive to sound movement has been described in the right parietal cortex⁵, which is dorsal to our scanned area. Furthermore, results from the visual domain indicate that the right hemisphere tends to be dominant for processing spatial information⁶, whereas the left human planum temporale may be specialized for speech processing⁷. Thus, in a motion context, certain types of sound pattern, as found in speech⁸, may introduce a left planum temporale bias into the described right dominance.

Previous studies have found evidence of

motion-sensitive auditory space maps in the midbrain⁹, patchy distributions of primary auditory-cortex neurons that are motion sensitive or motion selective¹⁰, and an area in the right human parietal cortex that is sensitive to sound movement⁵, which is presumably association cortex. Our results indicate that a missing neural link for this critical brain function may be a specialized area predominantly in the right posterior auditory cortex.

Frank Baumgart*,
Birgit Gaschler-Markefski*,
Marty G. Woldorff*†, **Hans-Jochen Heinze‡**,
Henning Scheich*

*Leibniz-Institute for Neurobiology,
Brenneckestrasse 6, 39118 Magdeburg, Germany

†Research Imaging Center,
University of Texas Health Science Center,
San Antonio, Texas 78284-6240, USA

‡Department for Neurology II,
Otto-von-Guericke-University,
Leipzigerstrasse 44, 39120 Magdeburg, Germany

1. Wagner, H., Kautz, D. & Poganiatz, I. *Trends Neurosci.* **20**, 583–588 (1997).

2. Knudsen, E. I. & Konishi, M. *Science* **200**, 795–797 (1978).

3. Scheich, H. et al. *Eur. J. Neurosci.* **10**, 803–809 (1998).
4. Woldorff, M. G. et al. *Hum. Brain Map.* **7**, 49–66 (1999).
5. Griffiths, T. D. et al. *Nature Neurosci.* **1**, 74–79 (1998).
6. Awh, E. & Jonides, J. in *The Attentive Brain* (ed. Parasuraman, R.) 353–381 (MIT Press, Cambridge, MA, 1998).
7. Galaburda, A. M., Le May, M., Kemper, T. L. & Geschwind, N. *Science* **199**, 852–856 (1978).
8. Belin, P. et al. *J. Cogn. Neurosci.* **10**, 536–540 (1998).
9. King, A. J. & Hutchings, M. E. *J. Neurophysiol.* **57**, 596–624 (1987).
10. Ahissar, M., Ahissar, E., Bergman, H. & Vaadia, E. *J. Neurophysiol.* **67**, 203–215 (1992).

Striped rabbits in Southeast Asia

The Annamite mountains of Laos and Vietnam have yielded several important mammalian discoveries¹. We have found a striped rabbit of the previously monospecific genus *Nesolagus*, extending its known range more than 1,500 km north from the island of Sumatra into mainland Southeast Asia. The Sumatran and mainland Annamite populations are morphologically similar, but genetic data indicate that they have been isolated for millions of years.

The first evidence of *Nesolagus* in the Annamite mountains came from freshly captured animals offered for sale in a food market in the rural town of Ban Lak, Laos (18°11' N, 104°58' E) between December 1995 and February 1996 (Fig. 1). The rabbits have since been found at other sites in the Annamite mountains of Vietnam.

The only previously known striped lagomorph was the critically endangered *Nesolagus netscheri*, a monotypic rabbit genus endemic to forest habitat within the Barisan

mountains and further north to Mount Leuser in Sumatra. There had been just one confirmed sighting of *N. netscheri* since 1916 and only around 15 museum specimens of the species exist, all of which were collected between 1880 and 1929 (ref. 2). In early 1998, *N. netscheri* was photographed by an automatic camera trap in the Kerinci Seblat National Park in Sumatra by Fauna and Flora International³. Externally, the Annamite rabbits closely resemble *N. netscheri*, having black or dark brown dorsal stripes, ferruginous rumps and short tails and ears (Fig. 2).

Morphological analysis of 30 characters (including pelage elements, external proportions, postcranial characters, tooth morphology and cranial morphometrics) in both Annamite and Sumatran rabbits reveals that there is a significant difference in the minimum interorbital distance as a percentage of condylobasal length (Annamite, range 19.0–23.6, mean (± s.d.) 20.5 (± 1.6), $n = 10$; Sumatran, 16.9–18.7, 17.5 (± 0.7), $n = 6$; $P < 0.001$, Wilcoxon–Mann–Whitney test).

Genetic analysis was undertaken of 653 continuous base pairs of the mitochondrial gene encoding 12S ribosomal RNA prepared from three museum specimens of *N. netscheri* and three rabbits from Laos. The results indicate that *N. netscheri* and the Annamite specimens (Genbank accession numbers AF176583 to AF176589) are sister taxa within a phylogenetic framework of lagomorph genera⁴. There is considerable divergence between them: the genetic distance (Kimura two-parameter) is 0.0552, which falls outside the range observed between other congeneric leporid species (0.0226–0.0427, mean 0.0324, from six species in two genera), but within that between different leporid genera (0.0469–0.1199, mean 0.0748, from seven genera). Assuming a steady rate of divergence over time at this gene⁵, the Sumatran and Annamite rabbits would have been diverging genetically for approximately 8 million years.

Despite an apparently high degree of conserved morphology, the large genetic divergence between these two taxa of striped rabbit indicates that they separated in the Pliocene epoch. In this region, glacial periods were cool and dry, characterized by expanding grassland, whereas forest cover increased during interglacial periods⁶. During glacial maxima, sea levels over the Sunda shelf were almost 150 metres lower than today, connecting Sumatra, Java and Borneo to the Asian mainland⁷. The ancestral *Nesolagus* may have been distributed over areas of this region at a time of lower sea level in the Pliocene, and its range would have been repeatedly dissected by periodic changes in both sea level and forest habitat.



Figure 2 Automatic camera-trap photographs of the striped rabbits. **a**, *Nesolagus netscheri* in Kerinci Seblat National Park in Sumatra (photograph from FFI). **b**, The striped rabbit from the Annamite mountains, photographed in the Pu Mat Nature Reserve, Vietnam (photograph from SFNC EC).

The existence of refugia for forest species during dry glacial periods has been postulated for both the Annamite mountains and Sumatra⁸. The contraction and expansion of ranges in and out of such allopatric refugia can generate divergence, with the sister genomes being protected by hybrid zones if they make contact, leading to speciation⁹. Our discovery of the rabbit in the Annamite mountains may provide insight into the factors governing current patterns of biodiversity in Southeast Asia, leading to its protection into the future.

Alison K. Surridge*, **Robert J. Timmins†‡**, **Godfrey M. Hewitt***, **Diana J. Bell***

*Population and Conservation Biology Sector, School of Biological Sciences, University of East Anglia, Norwich NR4 7TJ, UK
e-mail: a.surridge@uea.ac.uk
†Wildlife Conservation Society, PO Box 6712, Vientiane, Laos
‡Present address: 25 Cradley Road, Cradley Heath, West Midlands B64 6AG, UK



Figure 1 Map of Southeast Asia showing the Sunda islands of Sumatra, Borneo and Java. The site of the first discovery of a striped rabbit in the Annamites of central Laos is marked, as is the Kerinci Seblat National Park in Sumatra, where *Nesolagus netscheri* was photographed in the wild.

1. Gao, P. M. et al. *Anim. Conserv.* **1**, 61–68 (1998).
2. Flux, J. E. C. in *Rabbits, Hares and Pikas* (eds Chapman, J. A. & Flux, J. E. C.) 137–139 (IUCN, Gland, Switzerland, 1990).
3. Fauna and Flora International *Fauna Flora News* **9**, 3 (1998).
4. Halanaych, K. & Robinson, T. J. *J. Mol. Evol.* **48**, 369–379 (1999).
5. Allard, M. W. & Honeycutt R. L. *Mol. Biol. Evol.* **9**, 27–40 (1992).
6. van der Kaars, W. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **85**, 239–302 (1991).
7. Verstappen, H. T. in *Modern Quaternary Research in Southeast Asia* (eds Bartsa, G.-J. & Casparie, W. A.) 3–35 (Balkema, Rotterdam, 1975).
8. Brandon-Jones, D. *Biol. J. Linn. Soc.* **59**, 327–350 (1996).
9. Hewitt, G. M. *Biol. J. Linn. Soc.* **58**, 247–278 (1996).

A brilliant dissection of the brain

An Anatomy of Thought

by Ian Glynn

Weidenfeld and Nicolson: 1999. 456 pp. £25

Semir Zeki

Towards the end of his thrilling book, Ian Glynn tells us that scientists are generally dismissive of philosophy, partly at least because of the poor progress it has made compared to the 'hard' sciences. This is puzzling, if only because philosophy has largely dealt with the problem of knowledge, which is also a leading problem for those who study the brain, or at least its higher cognitive functions.

Like philosophy, brain science has not shed much light on the problem of the sources of knowledge that the brain has of the world. Indeed, it has perhaps not formulated it as clearly as have some philosophers, most notably Kant and Schopenhauer, the latter a special devotee of brain studies rarely acknowledged by brain scientists. But, with rare exceptions, philosophers have remained arrogantly indifferent to how the brain works. Brain science may not have made much progress with the problems of knowledge, but it has managed to learn a good deal about the brain.

Nowhere is this better exemplified than in comparing Glynn's book with those that have flooded the market in recent years from philosophers, and others in allied trades with an interest in the brain. Unlike theirs, Glynn's book has no ready, but questionable, solutions to the problems of the mind or of consciousness or their relationship to brain activity; neither is it assertive or arrogant. It is, instead, a biological treatise on the brain and it does not shy from discussing problems of the self, of consciousness, of creativity and of free will, because any serious consideration of the biology of the brain inevitably leads to these.

In preparing the reader for the delights to follow, Glynn begins with the problem of whether the laws of physics, as we know them, are sufficient to account for the emergence of life on Earth, or whether one should postulate another law or some vital force beyond. This really amounts to asking whether it is plausible that molecules characteristic of living things could be synthesized from non-living things. The balance of probabilities suggests that this is so, and that there is therefore no need to postulate any other agency. From that moment on, the diversity of life is determined by the slow and adventurous process of natural selection.

The highest achievement of that evolutionary process to date is the human brain, which has certain special qualities, among them the capacity to be conscious, to communicate through language, and to have



The brain and eye: the artist Bridget Riley juxtaposes colours in different sequences to "precipitate colour reactions" in the viewer. Glynn's book uses visual neuroscience as a test case.

insight and foresight. But the neurological mechanics of its knowledge and its deductive reasoning are unknown. These, no less than the possession of good colour vision or memory, have to be accounted for in terms of natural selection. Within such a biological context it is, for example, easy to question the supposition of Shadworth Hodgson that the mind is an epiphenomenon that cannot cause physical effects. If mental events cannot affect outcomes, and therefore have no survival values, why should we have evolved brains that make mental events possible?

This clears the ground here, and later in the book, to discuss altruistic behaviour and social relations and attitudes in terms of natural selection. Reading it from beginning to end, as one should, one emerges with a clear understanding of the biological issues that must be addressed to offer any satisfactory description of the brain, its functions and its functioning, its sources of knowledge and its

role in the ever-continuing process of evolution, not only physical but also social and intellectual.

The beginning of the book lays down the biological problems and traces the evolution of the human brain, and the end discusses issues of intentionality, consciousness, the mind and free will. In between, the reader is treated to a masterly account of the organization of the human brain, an account that even those well acquainted with the brain will learn a great deal from.

Glynn's erudition is astonishing and it makes these pages a hugely enjoyable intellectual journey, full of illuminating anecdotes that shed light not only on the discoveries but also on the scientists themselves: Darwin's failure, because of his seclusion at Down House, to read the work of Mendel, which was available at both the Royal Society and the Linnean Society; the insistence of the president of the Royal Society, when award-

FROM VISION: 50 YEARS OF BRITISH CREATIVITY (THAMES & HUDSON)

ing the Copley Medal to Darwin, that the council had “expressly omitted” *On the Origin of Species* from the “grounds of our award”; the experiments undertaken by Eduard Hitzig on his wife’s dressing-table because of the lack of facilities at the Physiological Institute in Berlin, and much else besides.

Perhaps my favourite tale is the account of how Pierre Marie, an eminent French neurologist, accused Déjerine, a French neurologist of even greater eminence, of doing his science as some others play roulette. Marie was challenged to a duel, whereupon he retracted the accusation — oh, how I wish one could still challenge scientists to duels! A similar accusation was once also made about Broca in relation to his discovery of a brain area critical for the production of articulate speech. In the end, the inspired Broca and Déjerine were proved right and the ever-cautious and careful Marie wrong.

I do not agree with everything that Glynn says. I am, for example, very sceptical about the phenomenon of blindsight, according to which subjects blinded because of lesions in their primary visual cortex are able to discriminate visual stimuli accurately without being conscious of having seen anything at all. I am also not at all certain that oscillations at 40 Hz are accepted as the mechanism that brings about the binding of the responses of cells in different brain areas — indeed, I think that most proponents of this doctrine have now abandoned it. But these disagreements do little to detract from the power of description and argument in this book.

Having given a comprehensive account of the human brain, Glynn finally turns to the problems of mind, consciousness and free will, where some philosophers come under the pickle-slicer. I am not sure that I agree with Glynn in his biological analysis of free will, but then I am not especially convinced of my own contrary arguments either. On consciousness, after having exposed the arguments with the usual clarity, he admits that “we don’t know yet what physical events are necessary and sufficient to cause specific conscious states — let alone why these events should have these causal powers” — *una grande lacuna*, as an Italian surgeon once lamented sadly when Glynn declared that he could not speak Italian. “Perhaps,” he says, “we are like our seventeenth-century forebears trying to understand nerve conduction before electricity was understood.”

At the end of the book, Glynn tells us with characteristic modesty that he cannot escape the charge made against Maimonides, that “no one should embark on a discussion of a problem unless he is capable of seeing it through to the end”. “But then,” Glynn adds, “I am not a sage.”

There has been *un’ altra grande lacuna*, one that Glynn has been able to fill admirably — a truly biological book about the most

perfect product of biological evolution: the human brain. It is this emphasis on biology that sets this book apart and makes it a gem that would have thrilled Kant and Schopenhauer. It should thrill modern readers even more. □

Semir Zeki is in the Wellcome Department of Cognitive Neurology, University College London, Gower Street, London WC1E 6BT, UK.

The problem with an evolutionary answer

The Dark Side of Man: Tracing the Origins of Male Violence

by Michael P. Ghiglieri

Perseus: 1999. 323 pp. \$26

R.C. Lewontin

Ever since the publication of Darwin’s *The Expression of the Emotions in Man and Animals* in 1872, people have been writing books about the evolution of human psychology and behaviour. After all, we are biological objects, the result of evolutionary forces just like other species, so why should human nature not be explicable by evolutionary transformations from the ancestors we have in common with other organisms?

The genre was given a major boost by Edward O. Wilson’s *Sociobiology* and *On Human Nature*. Ever since, there has been a stream of books explaining how this or that aspect of human nature is to be understood as a consequence of natural selection to maximize our contribution of genes to the next generation. Most of these have been rather more modest in their pretensions than Wilson’s original attempt to explain everything from aggression to xenophobia, and, not unexpectedly, there has been a disproportionate attention paid to the differences between men and women.

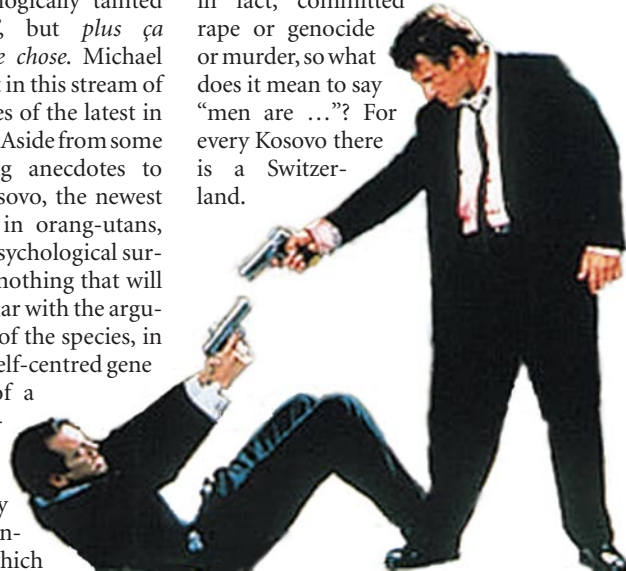
The term ‘sociobiology’ has recently been replaced by the less ideologically tainted ‘evolutionary psychology’, but *plus ça change, plus c’est la même chose*. Michael Ghiglieri’s book is the latest in this stream of always predictable syntheses of the latest in genes, apes and psychology. Aside from some updating of the shocking anecdotes to include the news from Kosovo, the newest observations of nastiness in orang-utans, and the latest outcome of psychological surveys on date rape, there is nothing that will interest those already familiar with the argument. ‘Man’, and the male of the species, in particular, is a violent and self-centred gene maximizer, the product of a deep biological core inherited from his ape-like ancestors. This core is overlain by a uniquely human self-conscious linguistic rationality, which

enables him to play out his violent inner nature in a rococo variety of culturally dependent forms.

Evolutionists seek to bring ‘human nature’ (often roped in to justify the inevitability of one social institution or another) under the explanatory mantle of evolutionary biology because they fear that their science is a failure if it cannot explain every aspect of the living world. But there is a confusion here between the undoubted fact that human life is a consequence of our properties as biological organisms and our ability to offer well founded and testable biological explanations of these properties.

Biologists, alas, have never been very good at distinguishing the ontological from the epistemological. There are two problems in the structure of evolutionary theory that cause particular difficulties in applying it to human individual and social behaviour, and they penetrate many domains of evolutionary studies. *The Dark Side of Man* epitomizes these problems and might make an excellent object of study for anyone trying to understand the deep problems of evolutionary explanation.

The first problem is in the tension between Darwinism as a variational theory of evolution and the typological description of a species’ nature. Ernst Mayr claims that Darwinism caused biologists to replace a typological view of organisms with a populational one. He ought to try reading *The Dark Side of Man* (or any of the other species of this literary genus). Human males are described as being by nature rapists, murderers, warriors and perpetrators of genocide, one chapter for each. I begin to doubt my own species identity, having never engaged in, or fantasized about, any of these activities whether drunk or sober, asleep or awake. And on reading the evidence discussed by Ghiglieri I discover that in the psychological studies only some men reported such activities or fantasies. Only a small fraction of men have, in fact, committed rape or genocide or murder, so what does it mean to say “men are ...”? For every Kosovo there is a Switzerland.



KOBAL COLLECTION

A typological description has been substituted for an immensely variable phenomenon, a typology that turns out to describe only a minority of the organisms. Biologists are always using typological descriptions to describe species that they believe have evolved, nevertheless, by a variational mechanism. They get away with it because the differences between species are usually very large compared with the intraspecific variation. But that is not true for human behaviour.

In this respect, human behaviour may be an evolutionary novelty, which raises the second big problem of evolutionary reconstruction. We depend on similarity for evolutionary stories and when faced with the threat of a novelty we either throw up our hands or, like Ghiglieri and other evolutionary psychologists, we assert a homology and bolster the claim by using the same words for different things. Orang-utans do not engage in 'genocide'. Only human beings can do that.

The real question about human social and individual behaviour is not why men belong to some type, but why they are so extraordinarily variable in time and space. □

R. C. Lewontin, of the Museum of Comparative Zoology at Harvard University, is at PO Box 274, Marlboro, Utah, USA.

The future of the doctors' drug culture

The Rise and Fall of Modern Medicine

by James Le Fanu

Little, Brown: 1999. 490 pp. £20

Ann Dally

James Le Fanu, general medical practitioner and distinguished medical journalist, has a fine overview of the changes in medicine of the past half century and he presents these as a spectacular success followed by recent "decline" and "fall". His view of medicine in 1945, before the great advances, is horrific and, in my view, exaggerated. Medicine then was not as barbaric as he depicts. Nevertheless, exaggeration heightens the impression of later "progress" and its contrast with the present situation, in which there is more discontent than ever before.

The medical profession, says Le Fanu, is preoccupied with "the shock of the new" and is full of disillusioned doctors treating "worried, well" patients obsessed with "trivial or non-existent threats to health". These days every newspaper proclaims such "dangers", which change constantly and send droves of "sufferers" to alternative practitioners. The bill for the UK National Health Service has doubled in a decade and things are getting worse.

Le Fanu describes beautifully what he



Scientific advances or quackery? Modern medicine has gone into decline and fall, argues Le Fanu.

regards as the 12 "definitive moments" that revolutionized medicine after the Second World War. The first was penicillin, which "transformed doctors' and the public's perceptions of medicine's possibilities" and "came to symbolise the almost limitless beneficent possibilities of science". Then came cortisone ('steroids'), and the mid-century watershed that produced both the statistical trials of streptomycin and PAS (*p*-aminosalicylic acid) and the proof that smoking caused lung cancer. (Later he argues that the development of statistical methods and surveys has been detrimental to medicine overall.)

Then followed open-heart surgery, chlorpromazine in psychiatry, hip replacements, kidney transplants, control of blood pressure to prevent strokes, the cure of childhood leukaemia (previously fatal), 'test-tube babies' and the discovery that peptic ulcers are caused by bacteria.

Following their discovery, some of the technologies were widely criticized for their inhumanity and for damaging patients for the sake of knowledge. Gradually, clinical research became more prestigious than clinical practice and introduced "something new and sinister" into medicine in which patients became "human guinea pigs". Much of it was "cruel, dangerous or purposeless", aimed solely at furthering the doctor's career by writing up the results in a scientific journal. The expansion was sophisticated, but where was it going?

Then triumphalism faltered. Problems grew and solutions became more elusive. The expensive Clinical Research Centre in Harrow, north London, was closed down after a decade. The *Medical Annual*, renowned for producing summaries of recent innovations, declined and expired.

Many new drugs were either useless or merely more expensive treatments than effective existing medication. The control on them was rudimentary. However, some were useful and one in particular was widely praised and prescribed, including to pregnant women. It was called thalidomide, which initiated a new era and much scepticism.

At this time, the public became aware that the interests of science did not always coincide with those of patients. The pharmaceutical industry turned to "blockbuster" drugs, such as Prozac and Viagra, and mergers occurred because companies now lacked ideas for new drugs, while patents on the older drugs were running out. Technology was out of step with major social trends. There were ever more sophisticated developments, the number of "scientific" tests done doubled in a decade, much effort was put into gathering unnecessary information or information that could have been obtained more easily, and many patients were made to suffer pain and misery, especially when dying.

Meanwhile, private medicine expanded and young doctors preferred private practice to clinical research. The number of dissatisfied patients increased, as did claims for medical negligence. Clinical science had lost its capacity to make substantial contributions to major medical problems, although these problems remained. Le Fanu blames not the technology itself, but "the intellectual and emotional immaturity of the medical profession".

Pride of place was taken by new versions of nature and nurture in the form of genetics and epidemiology. Le Fanu is contemptuous of both, not intrinsically, but because he believes that their potential for conquering

disease is small. "The New Genetics" focuses on abnormal genes that cause certain, mostly "staggeringly rare", diseases, which it aims to cure or prevent by changing these genes. In practice, genetics does little and is unlikely to do much more.

Epidemiology, "The Social Theory", emphasizes faulty diet and lifestyles as the cause of commoner diseases. Le Fanu questions the validity of much of this research, for which he is likely to attract criticism. He points out that none of the numerous allegations of environmental threats to health of the past 20 years has been validated. He even challenges the concept of poverty, or rather, modern Western poverty, as a cause of disease. He argues that current theories of epidemiology and genetics are manipulated to enhance the services of doctors and the profits of pharmaceutical companies, whose power and influence often corrupts doctors, and that the modern tendency of governments, or their medical officers, to recommend "safe" (from cancer) levels of various foods such as meat and alcohol is "indistinguishable from quackery".

Le Fanu ponders on the reasons for the unrealistic but pervasive belief in limitless possibilities for medical cures. He argues that the worship of "progress" and novelty in medicine must end if things are to improve, and that we should return to Sir William Osler's principles of observation and reasoning to know the true from the false — to prevent disease, to relieve suffering and to heal the sick through knowledge based on practical experience. Only then will doctors and patients be more satisfied. Unfortunately he makes no suggestions as to how this might be achieved, or even begun. If only one knew...

This is a challenging book. Some of the arguments may be oversimplified, but the message is clear and deserves careful consideration. □

Ann Dally is at the Wellcome Institute for the History of Medicine, 183 Euston Road, London NW1 2BE, UK.

Surfing the history of space

The Pearly Gates of Cyberspace

by Margaret Wertheim

Virago/W. W. Norton: 1999. 320 pp.

£14.99/\$24.95

Owen Gingerich

One has to admire the audacity of the concept of this book, subtitled *A History of Space from Dante to the Internet*. Science commentator Margaret Wertheim begins with the "soul space" of Dante's Christian geography, leads on to Giotto's emerging use of perspective in the Arena Chapel in Padua, describes the celestial space of Nicolaus Copernicus and Isaac Newton, and finally ends up with the Internet in its present and projected electronics world. Whether the book really holds together is another matter, but the juxtaposition of such varied views of space certainly provides a stimulating journey.

It is easy to develop a love-hate relationship with these essays, strung like pearls on a necklace. There is a fascinating sweep of ideas, with thoughtful research and novel insights. On the other hand, the book needed an editor who could have suppressed such a gratuitous description of Johannes Kepler as

"the weiner from Weil-der-Stadt", and who could have insisted on an index. The book could also have done with a copy editor who could spell "principal" and who could recognize that "millennia" and "phenomena" are plural nouns.

The chapter on relativistic space, which marches swiftly from Immanuel Kant to Edwin Hubble to Albert Einstein to Andrei Linde, actually says rather little about relativity, although it states, incredibly, that "without an understanding of special relativity, for instance, you would not have electric power coming efficiently to your home". It also imputes a curious thought process to Hubble — according to Wertheim, he simply imagined that the further away a nebula was, the faster it might be moving, and hence the more its spectrum might be redshifted. She concludes that the concept of the expanding Universe "was a brilliant imaginative leap, evidence that science does not proceed by logic alone". But surely the process went the other way around, starting with the observations of the high redshifts and correlating these with the faintness, and hence the farness of the nebulae. Hubble was convinced that the correlation existed, but was always somewhat ambivalent about the notion of the expanding Universe.

The final third of Wertheim's book deals with cyberspace, starting with a brief history of the explosive growth of the web, and including the burgeoning number of online fantasy worlds in which people take on elaborate alter egos. From chat rooms to MUDs (multi-user dungeons, laundered to multi-user domains), 'netizens' around the globe are engaging in addictive psychosocial experimentation. Wertheim claims that, after three centuries of physical materialism, "cyberspace helps to make explicit once more some of the *nonphysical* extensions of human beingness". It will, she says, challenge us to make a more nuanced conception both of ourselves and of the world around us.

Was cyber-utopia presaged by the London coffee-houses of Newton's day? Does modern cyberspace include an arena where the vilest sides of human behaviour can effloresce, a parallel to Dante's hell? Cyerpundit Hans Moravec has written that "wholesale resurrection may be possible through the use of immense simulators". Here, indeed, are the pearly gates of virtual reality and cyber immortality! These are some of the intriguing issues Wertheim addresses in her closing chapters.

"Like Copernicus," she writes, "we are privileged to witness the dawning of a new kind of space. What history will make of this space, appropriately enough, only time will tell." □

Owen Gingerich is at the Harvard-Smithsonian Center for Astrophysics, Cambridge, Massachusetts 02138, USA.

New in paperback

Life Itself: Exploring the Realm of the Living Cell

by Boyce Rensberger

Oxford University Press, £9.99, \$15.95

"Although some of the descriptions are rather lurid, what cannot be disputed is Rensberger's wonderment and enthusiasm because it literally leaps off the page. ... I am not entirely sure who *Life Itself* is aimed at. Presumably it is not intended to go up against the established cell biology tomes but it is hard to imagine a better way to convey to students the thrill of looking down a microscope at a living cell — and they get a pretty respectable introduction thrown in for good measure." Jeremy Hyams, *Nature* 386, 778 (1997)

The Calendar: Humanity's Epic Struggle to Determine a True and Accurate Year

by David Ewing Duncan

Avon, \$13.50

"The Calendar, although written in too

uncritical a style to be taken seriously as history, has a stronger chronological organization and relies on historical anecdote to keep the narrative moving. Consequently it is easier to read. It tells the story from the Stone Age to the Gregorian reform of the sixteenth century, and then steps nimbly on to conclude with the atomic clock. It offers a romp through history with calendars as the connecting thread." Jim Bennett, *Nature* 396, 328 (1998)

The Meaning of it All: Thoughts of a Citizen Scientist

by Richard P. Feynman

Penguin, £4.99

"He sets out to tackle science's relations with politics, religion and everyday society. If Feynman was a prophet, I suppose this was his sermon on divine uncertainty — the uncertainty that allows the scientific process to work. Knowledge can progress, says Feynman, only if people have open minds and test their ideas." Stephen Battersby, *Nature* 394, 144 (1998)

Images of Neptune's ring arcs obtained by a ground-based telescope

B. Sicardy*, F. Roddier†, C. Roddier†, E. Perozzi*‡, J. E. Graves†, O. Guyon†§ & M. J. Northcott†

* Observatoire de Paris, Institut Universitaire de France, Paris 6, DESPA, 92195 Meudon Cédex, France

† Institute for Astronomy, University of Hawaii, 2680 Woodlawn Drive, Honolulu, Hawaii 96822, USA

‡ Telespazio, Via Tiburtina 965, 00156 Roma, Italy

§ Ecole Normale Supérieure, 45 rue d'Ulm, 75230 Paris Cédex 05, France

Neptune has a collection of incomplete narrow rings, known as ring arcs, which should in isolation be destroyed by differential motion in a matter of months. Yet since first discovered¹ by stellar occultations in 1984, they appear to have persisted^{2–6}, perhaps through a gravitational resonance effect involving the satellite Galatea^{6–8}. Here we report ground-based observations of the ring arcs, obtained using an adaptive optics system. Our data, and those obtained using the Hubble Space Telescope (reported in a companion paper⁹), indicate that the ring arcs are near, but not within the resonance with Galatea, in contrast to what is predicted by some models.

Our data set consists of nine 600-s exposures taken at the Canada–France–Hawaii telescope in Hawaii between 11:44 and 13:15 UT on 6 July 1998, using a 1.72- μm filter. After standard flat-fielding and background subtraction, each image was deconvolved with an estimated point-spread function, clearly revealing some of the small inner satellites, namely Proteus, Larissa and Galatea, and more marginally, Despina. Neptune's centre was determined to within half a pixel (~ 0.017 arcsec) using both the motion of the satellites and the planet cloud cover. To enhance the signal-to-noise ratio, each pixel in individual images was rotated into Neptune's equatorial plane according to the orbital motion at the corresponding point, so as to show the neptunian system at a prescribed time, namely 12:29:39 UT (Earth time). A median filter was then applied when co-adding the nine images, while bright pixels near the planet were masked out. On this final image, Proteus, Larissa, Galatea and the arcs are detected, while Despina is lost in the glare of the planet (Fig. 1).

Small differences were found between the observed orbital phases of the satellites and those predicted from Voyager observations^{10,11}. Proteus, Larissa and Galatea are respectively 0.0 ± 0.3 , 1.65 ± 1.0 and 4.75 ± 1.7 degrees ahead of their expected longitude. However, the accumulated nominal errors since the Voyager observations in August 1989 are respectively 2.9, 5.2 and 8.1 degrees for the three satellites. Thus, none of these discrepancies in longitude are significant.

The arcs appear in Fig. 1 as a region of enhanced intensity along—but slightly outside—Galatea's orbit, near its maximum western elongation (right side of Neptune). Such a structure is clearly absent on the east side (left of Neptune). The arcs are embedded in a tenuous ring, Adams, which is about ten times fainter than the arcs⁴, and is thus undetectable in our images. To eliminate the background illumination, the intensity distribution on the east side was locally fitted by polynomials and then symmetrically subtracted from the west side. The arc intensity was then integrated over a width of 6 pixels (~ 0.2 arcsec) and plotted as a function of longitude (solid line in Fig. 2). This intensity is expressed in terms of an 'equivalent width', E_{arc} ; that is, the width of a perfect Lambert diffuser that would reflect sunlight at the distance of Neptune.

The dotted line in Fig. 2 shows the Voyager longitudinal arc profile, revealing the three main sub-arcs Liberté, Egalité and Fraternité⁶. Their position can be compared to the two possible solutions derived by Nicholson *et al.*³ for the mean

motion of the arcs n_{arc} , based on previous occultation detections and Voyager observations: $n_{\text{arc}} = 820.1194 \pm 0.0006 \text{ deg d}^{-1}$ and $n_{\text{arc}} = 820.1118 \pm 0.0006 \text{ deg d}^{-1}$, referred to as 'solution no. 1' and 'solution no. 2', respectively. These two solutions are easily distinguished in our data, since they yield a difference of 24.7 degrees in longitude on 6 July 1998. We find that the arcs are closer to the position predicted by solution 2. For instance, in Fig. 2, the Voyager

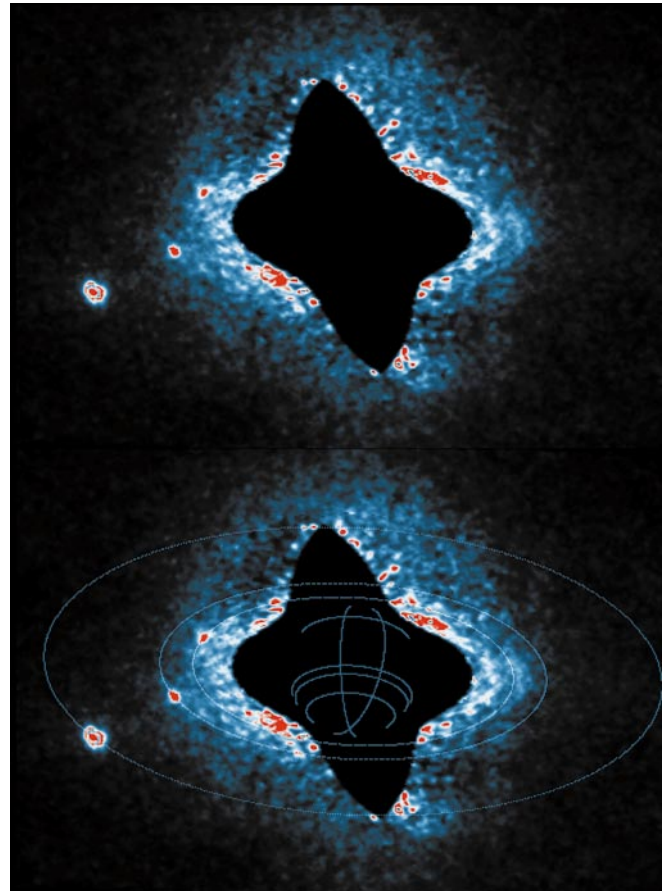


Figure 1 Images showing Neptune's ring arcs. Top, as observed; bottom, showing the orbits of the three moons (see below). The images were obtained by co-addition of nine 600-s exposures taken from the Canada–France–Hawaii 3.55-m f/36 telescope (CFHT), between 11:44 and 13:15 UT on 6 July 1998. Data were acquired through a narrow-band (120-nm) filter centred on a methane absorption band at 1.72 μm , thus minimizing the light scattered by Neptune. The image is 12 arcsec across, with a pixel size of 0.03425 arcsec. North is up and east is left, these directions being determined within 0.05 degrees by following the known motion of asteroid 554 Peraga with respect to stars during the same night. The CFHT was equipped with the 36-actuator adaptive optics (AO) camera Hokupa'a¹⁶, specifically developed at the University of Hawaii for astronomical applications¹⁶, and based on the concept of wave-front curvature sensing and compensation¹⁷. The AO technique compensates image degradation produced by the Earth's atmosphere. Optical aberrations are corrected at a kilohertz rate by means of a fast deformable mirror. Neptune itself (behind the black mask) was used as a guide source to sense the wave-front aberrations. Each image has first been rotated (see text) to show the system as it was at 12:29:39 UT, before being co-added to the others. On the bottom panel, ellipses have been drawn to show the computed orbits of Proteus, Larissa and Galatea, in order of decreasing distance to Neptune. The three satellites appear on the left side of Neptune as red spots, with a slight elongation due to their orbital motion during each of the 600-s exposures. The Adams ring arcs are seen at western elongation (right side of Neptune), just outside Galatea's orbit. The bright pixels between the planet and the arcs correspond the Le Verrier ring's orbit¹⁸. They are observed on only one side of the planet, and could be due to scattered light from the planet. New observations are needed to confirm this unexpected result.

dotted profile has been arbitrarily shifted by 4 degrees ahead of solution 2, in order to match the tallest peak (Fraternité) in our observed profile. We then note a mismatch of about 4 degrees in the positions of Liberté and Egalité, relative to Fraternité. Although it remains marginal in our data, this mismatch is confirmed by the recent Hubble Space Telescope (HST) observations⁹. Such change could be due to small libration motions of the arcs, as expected in the frame of the co-rotation model^{6,12}. Overall, the entire arc system is 5.5 ± 3 degrees ahead of solution 2, which yields an average arc mean motion of $n_{\text{arc}} = 820.1135 \pm 0.0009 \text{ deg d}^{-1}$ between August 1989 and July 1998. Considering our error bars, the difference between our value of n_{arc} and solution 2 remains marginal.

Among the models for arc confinement are co-rotation resonances with a nearby satellite. The equilateral Lagrange points L_4 or L_5 of a co-orbital satellite are examples of such stable sites¹³. More general co-rotation sites can be generated by an inclined satellite, outside its own orbit⁷. The Voyager and occultation data did show that the arcs lie very close to a co-rotation inclined resonance (CIR) with Galatea⁶. Due to its inclination I_G with respect to the local Laplace plane, Galatea's perturbing potential contains terms revolving at the pattern speed $n_{\text{cor}} = [(m_{\text{cor}} - 1)n_G + \dot{\Omega}_G]/m_{\text{cor}}$, where n_G and $\dot{\Omega}_G$ are Galatea's mean motion and nodal precession rate, respectively, and where the integer m_{cor} is the co-rotation wavenumber. The co-rotation occurs when $n_{\text{arc}} = n_{\text{cor}}$, and it can be shown that $2m_{\text{cor}}$ co-rotation sites are then created⁷.

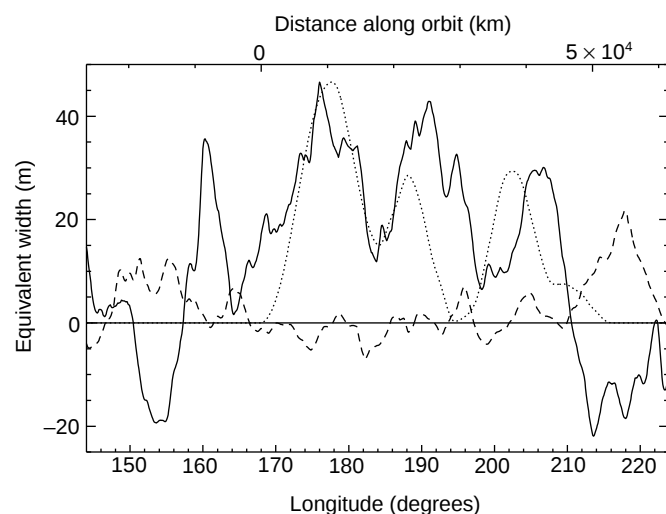


Figure 2 Analysis of the ring arcs. The thick solid line shows the equivalent width of the arcs derived from Fig. 1, as a function of longitude. The latter is counted from the ascending node of Neptune's equatorial plane on the 1950 Earth's Equator, for the epoch 8 August 1989, 12:00 Ephemeris Time at Neptune⁴, using a Neptune's spin pole direction at epoch (6 July 1998) of right ascension $\alpha_{1950} = 298.2712^\circ$, declination $\delta_{1950} = 43.4190^\circ$ (ref. 19). The orbital motion of the arcs is from left to right (that is, towards increasing longitudes). The peak at 160 degrees and the drop at 155 degrees are caused by noise. The dashed line shows the noise level in a region just beyond the arc. Note the increase of noise on each side of the arcs due to the light scattered by Neptune. The dotted line shows the Voyager profile convolved by the orbital motion of the arcs during the 600-s exposures, taking the equivalent width derived from low-phase-angle Voyager images⁴. The three main arcs are Liberté (right), Egalité (middle) and Fraternité (left). A fainter arc, Courage, is visible as a bump just ahead of Liberté in the dotted profile. The Voyager profile has been arbitrarily shifted to match the position of the trailing arc Fraternité, as explained in the text. The integrated H (1.65- μm) magnitude of the entire arc system shown here is 21 ± 0.5 . Note that the equivalent width that we obtain for the two brightest arcs Egalité and Fraternité (at a phase angle of 0.56 degrees) is comparable to the value derived from the clear filter images of Voyager (band-pass 0.31–0.56 μm) at low phase angle, ~ 8 degrees (ref. 4). Thus, within our error bars, we cannot detect any significant differences in the arc brightness between the visible and the near-infrared.

The wavenumber found for this CIR is $m_{\text{cor}} = 43$, which creates 86 co-rotation sites spanning each $360/2m_{\text{cor}} = 4.186$ degrees in longitude⁶. This resonance is referred to as the 42:43 CIR. From refs 6, 10 and 12, one can derive a maximum full width of each co-rotation site of $W_{\text{cor}} = 0.5 \pm 0.2 \text{ km}$. Consequently, if the co-rotation model is correct, the semi-major axis of the arcs should lie within $0.25 \pm 0.1 \text{ km}$ from the co-rotation radius for the arcs to be stably trapped. The dynamical, and to a lesser extent, the collisional, stability of this model have been successfully tested by numerical integrations^{8,12,14}.

As almost nine years elapsed since the Voyager observations, we derive a significantly improved value of $n_G = 839.66126 \pm 0.00052 \text{ deg d}^{-1}$, which agrees with the Voyager value¹⁰, 839.6598 ± 0.0025 , within the error bars. The parameter $\dot{\Omega}_G$ can be accurately calculated from the dynamical parameters of Neptune¹⁰, yielding $\dot{\Omega}_G = -0.714 \pm 0.002 \text{ deg d}^{-1}$. We finally obtain $n_{\text{cor}} = 820.11765 \pm 0.00052 \text{ deg d}^{-1}$, where most of the uncertainty comes from the uncertainty on Galatea's mean motion.

We detect a small but significant drift between the mean motion of the arcs and the pattern speed of the CIR, $n_{\text{arc}} - n_{\text{cor}} = -(4.2 \pm 1.0) \times 10^{-3} \text{ deg d}^{-1}$. Between August 1989 and July 1998, this drift results in an observed lag in longitude of 14 ± 3 degrees for the arcs, with respect to the CIR potential. This lag represents more than three times the azimuthal extension of one given site (4.186 degrees), thus calling into question the validity of the model of azimuthal confinement by the 42:43 CIR. This is confirmed by the fact that the value of $n_{\text{arc}} - n_{\text{cor}}$ represents a difference between the geometrical semi-major axis of the arcs and the radius of the 42:43 CIR of $a_{\text{arc}} - a_{\text{cor}} = 0.210 \pm 0.05 \text{ km}$, thus putting the arcs very near the separatrix between stable libration and unstable circulation. None of the direct numerical integrations performed so far^{8,12} have shown evidence for a drift $n_{\text{arc}} - n_{\text{cor}}$ once a particle is stably trapped into a co-rotation site. One way of accounting for a non-zero value of $n_{\text{arc}} - n_{\text{cor}}$ would be to change the value of $\dot{\Omega}_G$ ($= -0.714 \text{ deg d}^{-1}$) by 0.2 deg d^{-1} ; such a large correction would require an unrealistically large nearby satellite perturbing Galatea.

We note that the arcs should also be perturbed by a nearby Galatea outer Lindblad resonance^{6,7}. With our new orbital solutions, the 42:43 outer Lindblad resonance is now situated 1.86 km inside the arc geometrical radius, instead of 1.65 km as previously calculated⁶. With this new value, the radial structure of the arcs, and of the entire Adams ring, may still be dominated by Galatea, as predicted by Porco⁶, even though the azimuthal confinement of the arcs by the CIR now appears problematical.

An alternative solution is that the arcs are in fact trapped near a Lagrangian point L_4 or L_5 of a hypothetical neptunian satellite, as originally proposed by Lissauer¹³. For instance, a 6-km icy satellite (the detection limit of Voyager⁵) would create co-rotation sites with a maximum radial width of $\sim 0.6 \text{ km}$, similar to the co-rotation sites generated by Galatea, but with a longitudinal extension which could easily encompass the full azimuthal range of the arc system (~ 40 degrees). In this case, the mean motion of the arcs would simply reflect the mean motion of this unknown satellite, while being still perturbed by the 42:43 CIR with Galatea. But this would require an apparently fortuitous proximity of this undetected satellite to the 42:43 CIR with Galatea, and appears for the moment as a rather *ad hoc* explanation. $\square$

Received 15 March; accepted 30 June 1999.

- Hubbard, W. B., Brahic, A., Sicardy, B., Elicer, L. R. & Vilas, F. Occultation detection of a Neptunian ring-like arc. *Nature* **319**, 636–640 (1986).
- Sicardy, B., Roques, F. & Brahic, A. Neptune's rings, 1983–1989: Ground-based stellar occultations observations. *Icarus* **89**, 220–243 (1991).
- Nicholson, P. D., Mosqueira, I. & Matthews, K. Stellar occultation observations of Neptune's rings: 1984–1988. *Icarus* **113**, 295–330 (1995).
- Porco, C. C., Nicholson, P. D., Cuzzi, J. N., Lissauer, J. J. & Esposito, L. W. in *Neptune and Triton* (ed. Cruikshank, D. P.) 703–804 (Univ. Arizona Press, Tucson, 1995).
- Smith, B. A. et al. Voyager 2 at Neptune: imaging science results. *Science* **246**, 1422–1449 (1989).
- Porco, C. C. An explanation for Neptune's ring arcs. *Science* **253**, 995–1001 (1991).

7. Goldreich, P., Tremaine, S. & Borderies, N. Towards a theory for Neptune's arc rings. *Astron. J.* **92**, 490–494 (1986).
8. Horanyi, M. & Porco, C. C. Where exactly are the arcs of Neptune? *Icarus* **106**, 225–235 (1993).
9. Dumas, C., Terrile, R. J., Smith, B. A., Schneider, G. H. & Becklin, E. E. Stability of Neptune's ring arcs in question. *Nature* **400**, 733–735 (1999).
10. Owen, W. M. Jr, Vaughan, R. M. & Synnott, S. P. Orbits of the six new satellites of Neptune. *Astron. J.* **101**, 1511–1515 (1991).
11. Roddier, C. *et al.* Satellites and rings of Neptune. *IAU Circ.* No. 7051 (1998).
12. Foryta, D. W. & Sicardy, B. The dynamics of the Neptunian Adams ring's arcs. *Icarus* **123**, 129–167 (1996).
13. Lissauer, J. J. Shepherding model for Neptune's arc ring. *Nature* **318**, 544–545 (1985).
14. Hänninen, J. & Porco, C. Collisional simulations of Neptune's ring arcs. *Icarus* **126**, 1–27 (1997).
15. Graves, J. E., Northcott, M., Roddier, F., Roddier, C. & Close, L. First light for Hoku-pa'a, 36-element curvature optics system at UH. *Proc. SPIE* **3353**, 34–43 (1998).
16. Roddier, F., Northcott, M. & Graves, J. E. A simple low-order adaptive optics system for near-infrared applications. *Publ. Astron. Soc. Pacif.* **103**, 131–149 (1991).
17. Roddier, F. Curvature sensing and compensation: a new concept in adaptive optics. *Appl. Opt.* **27**, 1223–1225 (1988).
18. Roddier, C. *et al.* Rings of Neptune. *IAU Circ.* No. 7108 (1999).
19. Jacobson, R. A., Riedel, J. E. & Taylor, A. H. The orbits of Triton and Nereid from spacecraft and earthbased observations. *Astron. Astrophys.* **247**, 565–575 (1991).

Acknowledgements. The work of E.P. at DESPA is supported by the G. Colombo research fellowship of the European Space Agency. Adaptive optics observations of Neptune at the University of Hawaii were supported by NASA.

Correspondence and requests for materials should be addressed to B.S. (e-mail: Bruno.Sicardy@obspm.fr).

Stability of Neptune's ring arcs in question

Christophe Dumas*, Richard J. Terrile*, Bradford A. Smith†, Glenn Schneider‡ & E. E. Becklin§

* Jet Propulsion Laboratory, 4800 Oak Grove Drive, MS 183-501, Pasadena, California 91109-8099, USA

† University of Hawaii, Institute for Astronomy, 82-6012 Pu'uuhonua Road, Napo'opo'o, Hawaii 96704, USA

‡ University of Arizona, Steward Observatory, 933 N. Cherry Avenue, Tucson, Arizona 85721, USA

§ University of California at Los Angeles, Department of Physics and Astronomy, 405 Hilgard Avenue, Los Angeles, California 90095-1562, USA

Although all four of the gas-giant planets in the Solar System have ring systems, only Neptune exhibits 'ring arcs'—stable clumps of dust that are discontinuous from each other¹. Two basic mechanisms for confining the dust to these arcs have been proposed. The first² relies on orbital resonances with two shepherding satellites, while the second³ invokes a single satellite (later suggested to be Galatea⁴) to produce the observed ring arc structures. Here we report observations of the ring arcs and Galatea, which show that there is a mismatch between the locations of the arcs and the site of Galatea's co-rotation inclined resonance. This result calls into question Galatea's sole role in confining the arcs.

A new determination of the mean motion of the arcs was made possible by the Hubble Space Telescope (HST), using the Near Infrared Camera Multi-Object Spectrometer⁵ (NICMOS), which observed the system of Neptune's faint arcs on two occasions in 1998. To aid in the rejection of background light at the radial angular distance of the ring arcs (~ 3 arcsec) from the planet centre, a filter with a bandpass centred at $1.87\ \mu\text{m}$ wavelength was used. This wavelength corresponds to a strong absorption in the reflective spectrum of methane, one of the main constituents of Neptune's atmosphere. The projected azimuthal resolution was 3.2 degrees per resolution element (1.6 degrees per pixel) at the distance of the Adams ring, which corresponds to a resolution of 3,500 km along the ring. The first images of the arcs since the Voyager fly-by⁶ were obtained by NICMOS on 3 June 1998. The arcs were clearly detected at the maximum of eastern elongation (Fig. 1) during a single 832-s exposure. The orbital motion of the arcs smeared the image by 8 degrees along the azimuthal direction; this is about the full length of the longest arc, Fraternité. Nevertheless, the mean motion was

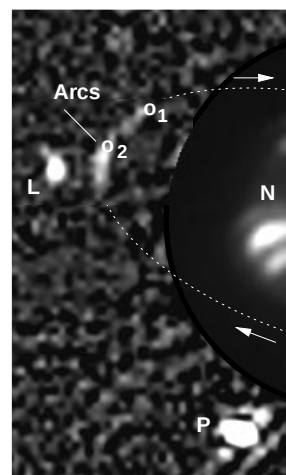


Figure 1 Initial image of Neptune's ring arcs obtained by the HST and its near-infrared camera NICMOS on 3 June 1998 at 19:40 UT. We used camera 2 (~ 0.076 arcsec per pixel) to image the vicinity of Neptune at $1.87\ \mu\text{m}$ while the planet's phase and aspect angles were respectively 1.9 and 63 degrees. Neptune's disk was positioned partly outside the field of view of the camera to reduce the scattered light coming from the planet. In this image, Neptune's contribution to the background has been modelled and removed. An image of Neptune's clouds has been superimposed for better clarity of the system geometry. The two circles (labelled 1 and 2) correspond to the positions of the middle-point of the trailing Egalité arc derived from the two possible solutions⁷ for the arcs' mean motion: $n_1 = 820.1194\ \text{deg d}^{-1}$ and $n_2 = 820.1118\ \text{deg d}^{-1}$. The letters L, P and N mark respectively the positions of Larissa, Proteus and Neptune, and the arrow shows the direction of motion of the ring arcs along their orbit. The measured location of the trailing Egalité arc fits better the position obtained with the mean-motion solution n_2 than n_1 .

unambiguously determined to be very close to the proposed value (or 'solution') of $820.1118\ \text{deg d}^{-1}$ (ref. 7). Before these observations, another solution ($820.1194\ \text{deg d}^{-1}$) was preferred because it supported better the action of the 42:43 co-rotation inclined resonance (CIR) with Galatea⁴.

We then modified our observational strategy to improve the image quality and spatial resolution for the two remaining HST orbits. Additional NICMOS observations were obtained on 20 and 22 October 1998. An image of the system of ring arcs is given in Fig. 2, as seen by an observer viewing from the ring-plane normal. The arcs extend over 40.7 ± 3.3 degrees and the brightest trailing arcs, Egalité and Fraternité, are clearly seen. The faintest leading arcs, Courage and Liberté, are just above the noise-level (3.5σ above background for Courage, the faintest arc). The centre of Egalité was used as a fiducial point to measure the locations of the ring arcs. The arcs were found to be 0.8 ± 1.5 degrees ahead of the ephemerides prediction employing a mean motion of $820.1118\ \text{deg d}^{-1}$. This measurement confirmed without ambiguity the results obtained during the June 1998 observations⁸. The new value for the arcs' mean motion is $820.1120 \pm 0.0004\ \text{deg d}^{-1}$. Similarly, we determined a revised mean motion for Galatea of $839.6617 \pm 0.0004\ \text{deg d}^{-1}$.

The brightness profile in backscattered light (Fig. 3) was obtained by integrating the flux, radially, over the photometric width of the arcs. Following the practice employed in earlier studies of Neptune's rings^{6,9}, we measured the equivalent width E_i seen by NICMOS in the optically thin case. Assuming that the optical depth of Egalité has not changed since 1989, its equivalent width, E_i , and geometric albedo, $p_{1.87\ \mu\text{m}}$, are 45 ± 5 m and 0.048 ± 0.005 , respectively. The latter is very close to that measured by Voyager at visible wavelength ($p_{0.5\ \mu\text{m}} = 0.055 \pm 0.004$), demonstrating that the colour of the material making up the ring arcs is similar to that in the rings of

Uranus¹⁰; that is, very dark with a relatively flat spectral reflectivity. Comparison between the HST brightness profile and the Voyager profile obtained at high phase angle (135°) shows that the apparent structure of the arcs has remained roughly the same since 1989. The main change concerns Liberté, which appears to be displaced from its position relative to Egalité by about 1.9 degrees in the leading direction. We assume that this shift cannot be solely caused by differences in the particle size distribution in Liberté, as the Voyager results showed similar brightness ratios among the three widest arcs at high and low phase angles.

Without stabilizing external forces, keplerian shear would spread the material of the ring arcs along the orbit of the Adams ring in a matter of a few months. The model of a single satellite responsible for both the radial and azimuthal confinement of the ring-arcs' material was first proposed in 1986 by Goldeich *et al.*³. This theory was supported by the work of Porco *et al.*^{4,11} who showed that Galatea, orbiting at the inner edge of the Adams ring on a slightly inclined trajectory (0.0544 degrees; ref. 12), was the most plausible body responsible for the long-lived stability of the arcs. The satellite

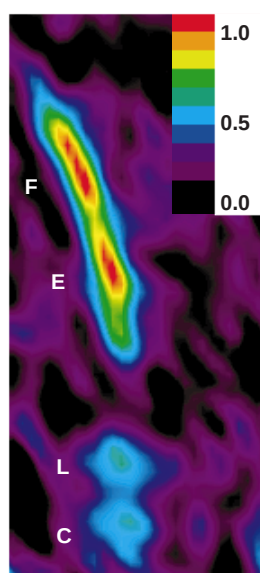


Figure 2 Higher-resolution false-colour image of the four ring arcs of Neptune obtained by HST/NICMOS at a wavelength of 1.87 μm . A total of 22 dithered images of 208 s each were recorded on 20 June 1998 (12:27 UT) and 22 October 1998 (17:40 UT). A prediction of 0.02 pixel (1.5 mas) in the co-registration of the images by cubic convolution interpolation was achieved by minimizing the difference in the diffraction pattern of Neptune in adjacent images. A model of the planet's scattered light was subtracted from the individual images in order to (1) recover a background level close to zero at the distance of the arcs from Neptune's centre, and (2) eliminate stray light produced by the secondary mirror support structure. The signal produced by the brightest arcs (Egalité/Fraternité) was at the noise level (0.065 μJy per pixel) in the individual images. The arcs were extracted from the individual images by taking into account their motion in the planetocentric reference system linked to Neptune. During the co-registration process, the data were resampled by a factor of two onto a finer grid by cubic convolution interpolation. The physical centre point of the system of satellites and arcs was determined by using the inner neptunian satellites (Proteus, Larissa, Despina and Galatea) present in our images to fit their individual trajectory whose pole solution has been known¹² since Voyager. The immediate benefit of the recentring, shifting and adding processes was to increase the resolution of our images to a point where the error of our measurement of the position of the arcs along the Adams ring was reduced to ± 1.5 degrees. The letters C, L, E and F indicate, from leading to trailing, the respective location of the arcs Courage, Liberté, Egalité and Fraternité. The colour has been scaled so the brightest arc has a value of unity.

would be responsible for the radial confinement of the dust through its 42:43 Linblad resonance (this mechanism was confirmed by Voyager observations of a 30-km radial distortion attributed to the Linblad resonance with Galatea). At the same time, its 42:43 CIR would prevent the arc material from spreading azimuthally. The relation between the motions of the CIR site and Galatea is defined as $m_c n_{\text{CIR}} = (m_c - 1)n_G + \dot{\Omega}$, where $\dot{\Omega}$ is Galatea's nodal precession rate, n_G the motion of Galatea, and m_c is the co-rotation resonance wavenumber. The location of the CIR produced by Galatea that falls closest to the arcs' semi-major axis is obtained for $m_c = 43$. Assuming that the value of $\dot{\Omega}$ determined by Voyager¹² remains unchanged—which is reasonable since Galatea's nodal precession rate is essentially determined by Neptune's gravitational harmonics—our new determination of the mean motions of Galatea and the ring arcs position the arcs 311 ± 41 m away from the middle of the CIR site. If we adopt a value for the half-width of the resonance site of 250 m (ref. 13), the mean semi-major axis of the arcs is located outside the region defined by the CIR with Galatea. This measurement calls into question the unique role of Galatea in confining azimuthally the material inside the arcs, leaving the field open for alternative models. We note that the wider (9 km; ref. 4) Linblad resonance remains a valid process to provide the arcs with the required energy to maintain their radial confinement.

Another possible scheme, as suggested originally by Lissauer², relies on two satellites instead of one. Although this model is based on the presence of an hypothetical moonlet co-orbiting with the Adams ring and responsible for trapping the dust between its L_4 and L_5 lagrangian points, it would solve the problem raised by the two trailing arcs Egalité and Fraternité that have widths well over the length (4.186 degrees) predicted by the 42:43 CIR with Galatea. Our

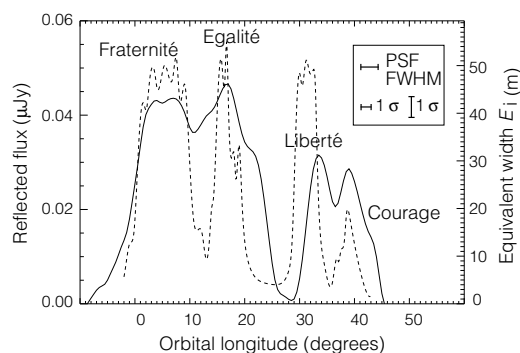


Figure 3 Brightness profile of the arcs along the azimuthal direction. The azimuthal profile of the arcs derived from Fig. 2 is shown here (solid line) both as the NICMOS radially integrated flux in μJy per azimuthal degree and as the equivalent width (E_1) in metres. Relative brightness of the arcs from Voyager (1989) observations, normalized to $E_1 = 52$ m (ref. 9) at Egalité, has been overplotted (dashed line) for comparison. We note that the Voyager profile was derived from an image recorded at a relatively high phase angle (135 degrees), where the observed brightness of the arcs is dominated by forward scattering from microscopic dust particles. However, within the measurement uncertainties, the same overall structure was seen in both low and high phase angle Voyager images. The 1σ uncertainties in both the NICMOS observed flux and azimuthal position are also indicated as well as the width (full-width at half-maximum, FWHM) of the NICMOS point-spread function (PSF) at 1.87 μm . The 1σ uncertainty in flux comes from the uncertainty in the determination of the residual background level introduced by the proximity of Neptune. The total integrated flux at 1.87 μm for all of the arcs is $27.0 \pm 0.27 \mu\text{Jy}$, or $M_{1.87 \mu\text{m}} = 21.2 \pm 0.1$ mag. The NICMOS calibration numbers used for filter Nic2/F187W are: conversion factor = 3.819 μJy per data number (DN) per s and zero-magnitude point = 828 Jy. The most apparent changes in the ring-arcs' morphology since 1989 are the widening of Egalité and the decrease in amplitude of Liberté. An effect induced by different scattering properties between high and low phase angles is not likely but cannot be entirely ruled out.

HST observations, along with the ground-based results presented elsewhere in this issue¹⁴, show that the new generation of large telescopes should allow us to collect additional measurements that will be valuable in constraining the theories of ring-arc formation around Neptune. □

Received 18 March; accepted 6 July 1999.

- Hubbard, W. B. *et al.* Occultation detection of a neptunian ring-like arc. *Nature* **319**, 636–640 (1986).
- Lissauer, J. J. Shepherding model for Neptune's arc ring. *Nature* **318**, 544–545 (1985).
- Goldreich, P., Tremaine, S. & Borderies, N. Towards a theory for Neptune's arc rings. *Astron. J.* **92**, 490–494 (1986).
- Porco, C. C. An explanation for Neptune's ring arcs. *Science* **253**, 995–1001 (1991).
- Thompson, R. L., Rieke, M., Schneider, G., Hines, D. C. & Corbin, M. Initial on-orbit performance of NICMOS. *Astrophys. J.* **492**, L95–L97 (1998).
- Smith, B. A. *et al.* Voyager 2 at Neptune: Imaging science results. *Science* **246**, 1422–1449 (1989).
- Nicholson, P. D., Mosquera, I. & Matthews, K. Stellar occultation observations of Neptune's rings. *Icarus* **113**, 295–330 (1995).
- Terrile, R. J. *et al.* First infrared imaging of the Neptune ring arcs: HST/NICMOS results. *Bull. Am. Astron. Soc.* **30**, 1044 (1998).
- Porco, C. C., Nicholson, P. D., Cuzzi, J. F., Lissauer, J. J. & Esposito, L. W. in *Neptune and Triton* (ed. Cruikshank, D. P.) 703–806 (Univ. Arizona Press, Tucson, 1985).
- Elliot, J. L. & Nicholson, P. D. in *Planetary Rings* (eds Greenberg, R. & Brahic, A.) 25–72 (Univ. Arizona Press, Tucson, 1984).
- Horany, M. & Porco, C. C. Where exactly are the arcs of Neptune? *Icarus* **106**, 525–535 (1993).
- Owen, W. M., Vaughan, R. M. & Synnott, S. P. Orbits of the six new satellites of Neptune. *Astron. J.* **101**, 1511–1515 (1991).
- Foryta, D. W. & Sicardy, B. The dynamics of the neptunian Adams ring's arcs. *Icarus* **123**, 129–167 (1996).
- Sicardy, B. *et al.* Images of Neptune's ring arcs obtained by a ground-based telescope. *Nature* **400**, 731–733 (1999).

Acknowledgements. We thank R. Thompson and M. Rieke from the NICMOS IDT team as well as P. Nicholson for calculating the predicted offsets of the arcs for 20 and 22 October 1998, and R. Jacobson for providing us with the coordinates of the normal of the orbital planes of the neptunian inner satellites.

Correspondence and requests for materials should be addressed to C.D. (e-mail: Christophe.Dumas@pl-nasa.gov).

Linking insulator-to-metal transitions at zero and finite magnetic fields

Y. Hanein*, N. Nenadovic*, D. Shahr*, Hadas Shtrikman*, J. Yoon†, C. C. Li† & D. C. Tsui†

* Department of Condensed Matter Physics, Weizmann Institute, Rehovot 76100, Israel

† Department of Electrical Engineering, Princeton University, Princeton, New Jersey 08544, USA

For many years, it was widely accepted¹ that electrons confined to two dimensions would adopt an insulating ground state at zero temperature and in zero magnetic field. Application of a strong perpendicular magnetic field changes this picture, resulting^{2,3} in a transition from the insulating phase to a metallic quantum Hall state. Unexpectedly, an insulator-to-metal transition was recently observed⁴ in high-quality two-dimensional systems at zero magnetic field on changing the charge carrier density. The mechanism underlying this transition remains unknown^{5–9}. Here we investigate the magnetic-field-driven transition in a two-dimensional gallium arsenide system, which also exhibits^{10–12} the poorly understood zero-field transition. We find that, on increasing the carrier density, the critical magnetic field needed to produce an insulator-to-metal transition decreases continuously and becomes zero at the carrier density appropriate to the zero-field transition. Our results suggest that both the finite- and zero-magnetic field transitions share a common physical origin.

In Fig. 1a we plot the resistivity (ρ) of one of our samples as a function of magnetic field (B) at several temperatures, with the density of holes (p) held fixed. The sample is a p-type, inverted semiconductor–insulator–semiconductor (ISIS) sample¹³ grown on the (311)A GaAs substrate¹⁴ with silicon as a p-type dopant. At $B = 0$ the system is insulating, as indicated by a rapidly increasing ρ

as the temperature (T) approaches zero. The insulating behaviour is maintained for $B < B_c^L$. For $B > B_c^L$, a quantum-Hall (QH) state ($\nu = 1$, where ν is the Landau-level filling factor) is observed with ρ tending to zero as T is lowered. We identify B_c^L , the point where the temperature coefficient of resistivity (TCR) changes its sign, with the critical point of the insulator-to-QH transition². At still higher B , beyond the $\nu = 1$ QH state, the system turns insulating and a second T -independent transition point is seen at B_c^H . B_c^H therefore marks the critical B of the QH-to-insulator transition¹⁵. Following the path set by earlier studies^{10,16,17} we focus, for now, on the low- B transition and follow, in Fig. 1b–e, the evolution of the critical point B_c^L as we increase p .

Data obtained from the same sample at successive increases of the hole density are shown in Fig. 1b–e. Similarly to Fig. 1a, a low- B transition point is evident in Fig. 1b, but occurs at a lower B . This trend continues in Fig. 1c until finally, in Fig. 1d, the crossing point disappears. In addition to the shift of B_c^L towards lower B , the higher p causes more QH states to become experimentally observable.

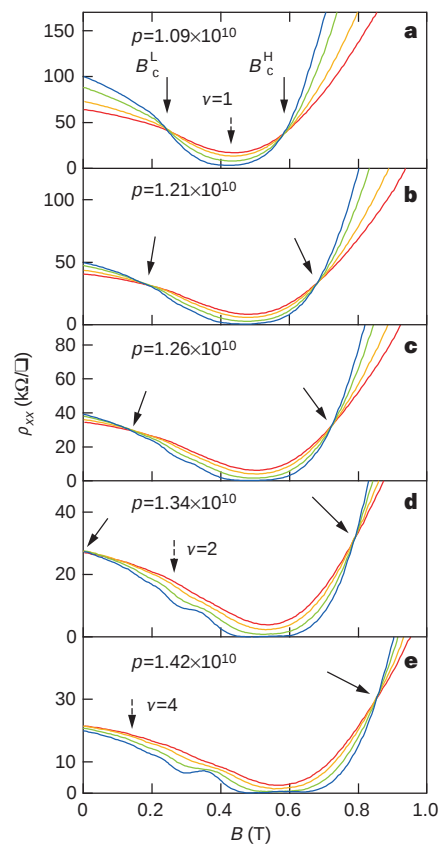


Figure 1 Isotherms of the magnetoresistance data of sample H324Bc at various hole densities. In ISIS samples, the hole carrier-density (p) is varied by applying voltage (V_g) to the grown back-gate of the ISIS. The ohmic contacts to the two-dimensional layer have to be alloyed very carefully to avoid an electrical short between this layer and the underlying gate. As a result, the use of an ISIS structure is accompanied by complications in realizing reliable ohmic contacts at extremely low p s and low T s. In most of our devices, only some of the alloyed contacts survive at low T and low p . For the sample in this study, we do not have Hall measurements compatible with the set of measurements we present here. The sample was wet-etched to the shape of a standard 'Hall bar' (60 × 120 μm) and the measurements were done in a dilution refrigerator with a base T of 57 mK, using the a.c. lock-in technique with an excitation current of 1 nA flowing in the [011] direction. The traces are colour coded with blue being low, and red being high, temperature: values of 63, 122, 177 and 217 mK were used at each p value. The solid arrows indicate the position of the crossing points, and the dashed arrows the position of the $\nu = 1, 2$ and 4 filling factors. Resistivity along the current direction (ρ_{xx}) was measured in units of $k\Omega$ per square.

Hence the transition is no longer from an insulator to the $\nu = 1$ QH state, but involves the $\nu = 2$ (and higher, Fig. 1c) state. Eventually, the metallic behaviour at the high- B side of the transition is too weak to be experimentally identified with a QH state.

Along with the shift in B_c^L and the evolution of the QH states, we notice in Fig. 1a–c that the insulating behaviour at $B = 0$ becomes weaker until, in Fig. 1d, ρ at $B = 0$ is T -independent. In Fig. 1e, the system has crossed over into its metallic phase and no low- B transition, or T -independent point, is seen; this implies that the hole density shown in Fig. 1d ($p = 1.34 \times 10^{10} \text{ cm}^{-2}$) is the critical density of the metal–insulator transition (MIT) at $B = 0$. This $B = 0$ transition is the MIT in two dimensions first reported by Kravchenko *et al.* for Si samples⁴. Our main result can now be stated: on increasing p , B_c^L gradually tends to lower B s, eventually converging to the $B = 0$ MIT which, for this sample, takes place at $p = 1.34 \times 10^{10} \text{ cm}^{-2}$.

To complement our p -dependence study of the B -driven transitions, we next focus on the effect of a perpendicular magnetic field on the p -driven transition. Our new starting point is the more conventional experimental demonstration of the $B = 0$ MIT in two dimensions. In Fig. 2a we plot ρ as a function of p at several T s and at $B = 0$. A T -independent crossing point is seen here as well (at $p_c = 1.34 \times 10^{10} \text{ cm}^{-2}$), marking the transition from insulating behaviour for $p < p_c$ to metallic behaviour for $p > p_c$. We then repeat the measurement of Fig. 2a at different values of B (Fig. 2b–e). In Fig. 2b and c, p_c shifts to a lower value, a trend which reverses for $B > 0.35$ T (Fig. 2d and e). This trend reversal of $p_c(B)$ is accompanied by the development of non-monotonic dependence of

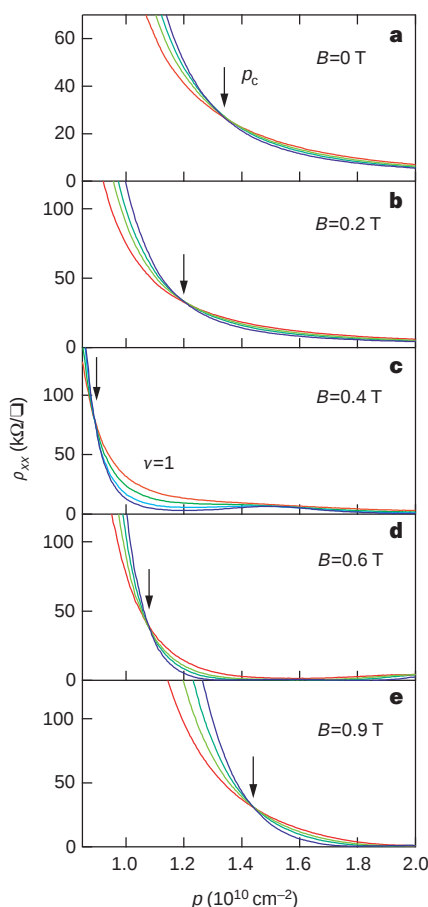


Figure 2 Isotherms of the resistance versus p of sample H324Bc at various magnetic fields. The traces are colour coded with blue being low T , and red being high T . **a**, $B = 0$ T at $T = 57, 120, 160$ and 214 mK. **b**, $B = 0.2$ T at $T = 58, 121, 160$ and 217 mK. **c**, $B = 0.4$ T at $T = 60, 98, 148$ and 212 mK. **d**, $B = 0.6$ T at $T = 59, 122, 161$ and 217 mK. **e**, $B = 0.9$ T at $T = 59, 122, 161$ and 217 mK.

ρ on p , which is a precursor to the QH effect. We now combine these $p_c(B)$ results with the $B_c(p)$ of Fig. 1, to plot a comprehensive phase diagram of our system in the B – p plane.

The phase diagram obtained from our data is shown in Fig. 3, where we plot the B and p coordinates of each of the transitions. Separate symbols are given to B_c^L and B_c^H , defined in Fig. 1, and to p_c from Fig. 2. Several points emerge from inspecting the resulting phase diagram. First, we note that the results obtained from the two data sets (fixed p and fixed B measurements) are mutually consistent. Second, for fixed p s in the range $(0.88\text{--}1.33) \times 10^{10} \text{ cm}^{-2}$, the low- B insulating phase first turns metallic and then reappears at high B . This re-entrant nature of the insulating phase is clearly reflected in the B traces of Fig. 1a–c. And third, the low- B region of the phase boundary reiterates the main result of our work and clearly depicts the continuous evolution of the transition from the high- B MIT to the $B = 0$ MIT. The relation between the transition at finite B and the MIT transition at $B = 0$ suggests that similar processes govern the transport for both transitions¹⁸.

Figure 3 also includes the high- B side of the phase diagram (B_c^H). In fact, the low- and high- B regimes are smoothly connected to form a single phase-boundary line. It is common practice to describe the finite- B transitions in the language of quantum phase transitions¹⁹. If we assume that the phase-boundary line of Fig. 3 comprises a set of quantum critical points, it is possible that universal features should be observed in its vicinity. To test this proposition we examine the value of ρ at the transition points, ρ_c . In Fig. 4a we plot ρ_c of our transitions as a function of B . Overall, ρ_c is not constant, its value changing by almost a factor of 4 over our B range. However, at very low as well as at very high B , ρ_c approaches a value close to h/e^2 , the quantum unit of resistance. Although for our sample, at $B = 0$, ρ_c is close to h/e^2 , we point out that the value of ρ_c at $B = 0$ obtained for other samples varies by an order of magnitude, between $0.4h/e^2$ and $4h/e^2$ (ref. 20).

Because of the clear separation of the B – p plane into conducting and non-conducting regions, it is tempting to consider them as two unique phases for the entire B range. However, the nature of these phases may not be the same for all B values. Considering the metallic phase: at very low p , the metallic-like behaviour is due to the formation of the $\nu = 1$ QH state. At higher p s, the metallic side of the transition appears to be due to a higher-order QH state, with $\nu = 2$ or 4. However, ρ in these higher-order states does not vanish, and a clear determination of the nature of the metallic conduction is uncertain. In fact, due to the higher disorder at low p , a fully developed QH state is not expected at any experimentally reasonable T . The proximity to a MIT point is also a contributing factor to the ambiguity in determining the properties of the metallic region. Our data therefore preclude us from ruling out the possibility of the existence of an additional transition separating the metallic phase at

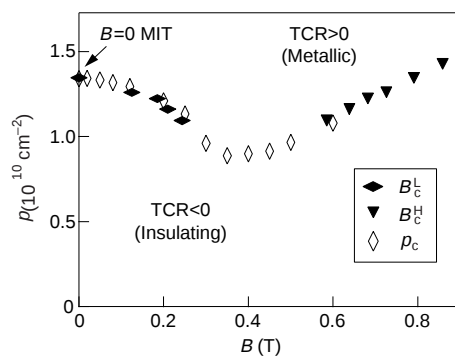


Figure 3 B and p coordinates of the crossing points. The latter are derived from data similar to those of Figs 1a–e and 2a–e. Open diamonds denote p_c , filled diamonds denote B_c^L and filled triangles denote B_c^H . The arrow points to the $B = 0$ metal–insulator transition (MIT). TCR is the temperature coefficient of resistivity.

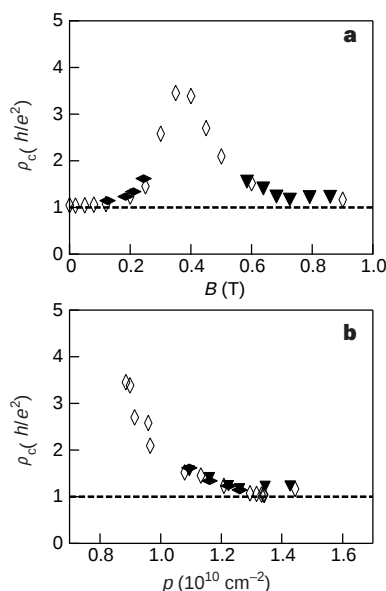


Figure 4 The value of ρ at the crossing points. The latter are derived from data similar to those of Figs 1a–f and 2a–e. **a**, ρ_c plotted as a function of B . Open diamonds denote ρ_c from p_c data, filled diamonds denote B_c^L data and filled triangles denote B_c^H data. **b**, ρ_c plotted as a function of ρ ; symbols as **a**.

low B and the QH phase at higher B . We note that experimental data^{21,22} suggest that in the limit of zero temperature a metallic phase does not exist at any finite B . Hence, the nature of the metallic behaviour for the entire B range remains open for future investigation.

We have shown that B_c^L , B_c^H and p_c define a common phase-boundary line in the B – p plane, and that at intermediate B s along this phase boundary ρ_c significantly deviates from h/e^2 , its value near $B = 0$ and at high B . It is instructive to consider the dependence of ρ_c on p , rather than on B . In Fig. 1a–c we can readily see the general trend: the ρ_c of the low- B and high- B transitions at fixed p are very close to each other. To test this result for our entire range of p we plot, in Fig. 4b, ρ_c versus p obtained from our data. The two transitions are seen to have collapsed onto a single curve for our entire range of p , indicating that, for a given carrier density, the ρ_c of the low- B and the high- B transitions is the same. This supports the notion of symmetry between these transitions²³.

A possible relation between different transitions in two-dimensional systems was noted by Jiang *et al.*², who pointed out the similarities between the insulator-to-QH state and the insulator-to-superconductor transitions; the theoretical basis for such similarity has been reported^{3,18,24,25}. In our work we found a relation between the finite- B conductivity transitions and the metal–insulator transition at $B = 0$. As all transitions are measured in the same two-dimensional system, we were able to continuously transform one to the other. This raises the possibility that both transitions share a common physical origin. □

Received 22 January; accepted 17 June 1999.

1. Abrahams, E., Anderson, P. W., Licciardello, D. C. & Ramakrishnan, T. V. Scaling theory of localization: Absence of quantum diffusion in two dimensions. *Phys. Rev. Lett.* **42**, 673–676 (1979).
2. Jiang, H. W., Johnson, C. E., Wang, K. L. & Hannahs, S. T. Observation of magnetic-field-induced delocalization-transition from Anderson insulator to quantum Hall conductor. *Phys. Rev. Lett.* **71**, 1439–1442 (1993).
3. Kivelson, S., Lee, D. H. & Zhang, S. C. Global phase diagram in the quantum Hall effect. *Phys. Rev. B* **46**, 2223–2238 (1992).
4. Kravchenko, S. V., Kravchenko, G. V., Furneaux, J. E., Pudalov, V. M. & D'Iorio, M. Possible metal-insulator transition at $B = 0$ in two dimensions. *Phys. Rev. B* **50**, 8039–8042 (1994).
5. Finkelstein, A. M. Influence of Coulomb interaction on the properties of disordered metals. *JETP* **57**, 97–108 (1983).
6. Dobrosavljevic, V., Abrahams, E., Miranda, E. & Chakravarty, S. Scaling theory of two-dimensional metal-insulator transitions. *Phys. Rev. Lett.* **79**, 455–458 (1997).
7. Altshuler, B. L. & Maslov, D. L. Theory of metal-insulator transitions in gated semiconductors. *Phys. Rev. Lett.* **82**, 145–148 (1999).

8. Phillips, P., Wan, Y. L., Martin, I., Knys, S. & Dalidovich, D. Superconductivity in a two-dimensional electron gas. *Nature* **395**, 253–257 (1998).
9. He, S. & Xie, X. C. New liquid phase and metal-insulator transition in Si-MOSFETs. *Phys. Rev. Lett.* **80**, 3324–3327 (1998).
10. Simmons, M. Y. *et al.* Metal-insulator transition at $B = 0$ in an ultra-low density two dimensional hole gas. *Physica B* **251**, 705–709 (1998).
11. Simmons, M. Y. *et al.* Metal-insulator transition at $B = 0$ in a dilute two dimensional GaAs-AlGaAs hole gas. *Phys. Rev. Lett.* **80**, 1292–1295 (1998).
12. Hanein, Y. *et al.* The metalliclike conductivity of a two-dimensional hole system. *Phys. Rev. Lett.* **80**, 1288–1291 (1998).
13. Meirav, U., Heiblum, M. & Stern, F. High-mobility variable-density two-dimensional electron gas in inverted GaAs-AlGaAs heterojunctions. *Appl. Phys. Lett.* **52**, 1268–1270 (1988).
14. Hanein, Y., Shtrikman, H. & Meirav, U. Very low density two-dimensional hole gas in an inverted GaAs/AlAs interface. *Appl. Phys. Lett.* **70**, 1426–1428 (1997).
15. Shahar, D., Tsui, D. C., Shayegan, M., Bhatt, R. N. & Cunningham, J. E. Universal conductivity at the quantum Hall liquid to insulator transition. *Phys. Rev. Lett.* **74**, 4511–4514 (1995).
16. Shashkin, A. A., Dolgoplov, V. T. & Kravchenko, G. V. Insulating phases in a two-dimensional electron system of high mobility Si MOSFET's. *Phys. Rev. B* **49**, 14486–14495 (1994).
17. Dultz, S. C., Jiang, H. W. & Schaff, W. J. Absence of floating delocalization states in a two-dimensional hole gas. *Phys. Rev. B* **58**, R7532–R7535 (1998).
18. Shimshoni, E., Auerbach, A. & Kapitulin, A. Transport through quantum melts. *Phys. Rev. Lett.* **80**, 3352–3355 (1998).
19. Sondhi, S. L., Girvin, S. M., Carini, J. P. & Shahar, D. Continuous quantum phase transitions. *Rev. Mod. Phys.* **69**, 315–333 (1997).
20. Hanein, Y., Shahar, D., Yoon, J., Li, C. C., Tsui, D. C. & Shtrikman, H. Observation of the metal-insulator transition in two-dimensional n-type GaAs. *Phys. Rev. B* **58**, R13338–R13340 (1998).
21. Simonian, D., Kravchenko, S. V., Sarachik, M. P. & Pudalov, V. M. Magnetic field suppression of the conducting phase in two dimensions. *Phys. Rev. Lett.* **79**, 2340–2307 (1997).
22. Pudalov, V. M., Bruthaler, G., Prinz, A. & Bauer, G. Instability of the two-dimensional metallic phase to a parallel magnetic field. *JETP Lett.* **65**, 932–937 (1997).
23. Hilke, M. *et al.* Symmetry in the insulator-Hall-insulator transitions observed in a Ge/SiGe quantum well. *Phys. Rev. B* **56**, 15545–15548 (1997).
24. Fisher, M. P. A. & Lee, D. H. Anyon superconductivity and the fractional quantum Hall effect. *Phys. Rev. Lett.* **63**, 903–906 (1989).
25. Fisher, M. P. A., Grinstein, G. & Girvin, S. M. Presence of quantum diffusion in two dimensions: Universal resistance at the superconductor-insulator transition. *Phys. Rev. Lett.* **64**, 587–590 (1990).

Acknowledgements. We thank E. Shinshoni, M. Hilke and A. Yacoby for discussions. This work was supported by the NSF, the US-Israel BSF and by the Israeli Ministry of Science and the Arts.

Correspondence and requests for materials should be addressed to Y.H. (e-mail: hanin@wis.weizmann.ac.il).

Observation of short-range critical wetting

D. Ross, D. Bonn & J. Meunier

Laboratoire de Physique Statistique de l'ENS, UMR 8550 du CNRS, associé aux universités Paris VI et Paris VII 24 rue Lhomond, 75231 Paris cedex 05, France

Mean-field theory correctly predicts the critical behaviour of systems close to a phase transition, provided that fluctuations can be neglected. Fluctuations, however, become important if the dimensionality of the system is lower than a certain upper critical dimension. For such systems, it is necessary to use renormalization-group methods to describe the critical behaviour. Investigation of three-dimensional systems in which the upper critical dimension is also three can therefore provide a probe of the way in which mean-field theory breaks down when fluctuations become important^{1–12}. An important example is the critical wetting transition that is predicted^{1,2} to occur in systems in which long-range forces are negligible, involving a continuous and reversible increase in the thickness of an adsorbed film. Here we present experimental observations of the short-range wetting transition close to the critical point in methanol–alkane binary liquid mixtures. We observe second-order, critical wetting for nonane (as characterized by the surface specific-heat exponent). The measured value is consistent with the predictions of mean-field theory, but disagrees strongly with renormalization-group calculations, which predict^{1–4} non-universal behaviour for this transition. The reasons for the apparent failure of the renormalization-group approach remain unclear; further experiments are needed to investigate the effects of fluctuations in more detail.

When a liquid droplet is put onto a surface, three different

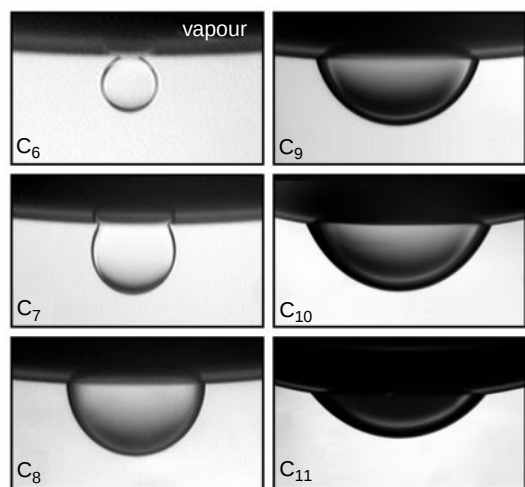


Figure 1 Photographs of methanol droplets at the liquid–vapour interfaces of the different *n*-alkanes at $T \approx 20^\circ\text{C}$ (for C_6 , $T \approx 3^\circ\text{C}$, as the wetting transition has already taken place at room temperature). The different *n*-alkanes, C_nH_{2n+2} , are denoted as C_n . The alkanes used were of purity $>99\%$, and the methanol purity was $>99.9\%$. For the photographs and the contact-angle measurements (Figs 2, 4), the methanol/alkane mixtures were sealed inside a glass ampoule which was submersed in a temperature-controlled water bath.

situations (distinguishable by the contact angle, θ) may result. First, if the contact angle is zero, the droplet spreads across the surface, a situation referred to as complete wetting. Second, for a contact angle of 180° , a layer of vapour intrudes between the droplet and the surface in a situation referred to as complete drying. Third, for contact angles between zero and 180° , the situation is called partial wetting. A wetting (or drying) transition is a surface phase transition from partial to complete wetting (drying). The wetting transition is generally first-order (discontinuous), implying a discontinuity in the first derivative of the surface free energy.

We show here that the wetting properties of methanol on the liquid–vapour interface of *n*-alkanes (the substrate) are unusual in that by changing the alkane chain length, the crossover from wetting to drying can be studied (Fig. 1). For each of the alkanes a wetting (drying) transition can be found where the contact angle goes to zero (180°) as the temperature is varied (Fig. 2). The ratio, T_w/T_c , of the wetting transition temperature to the critical temperature (at which the two liquids become miscible) varies from about 0.9 for undecane (C_{11}) to nearly 1 for nonane (C_9).

Measurements of the equilibrium thickness of the layer of methanol at the alkane–vapour interface (Fig. 3) show a marked difference between the wetting transitions in these two (usual) systems. For undecane, the usual first-order wetting transition is observed, manifested as a discontinuous jump in the film thickness from a microscopic value to a thick (gravitationally limited) value of about $100\text{ \AA}$. The surprising result is that the increase of the film thickness of methanol on nonane is continuous and reversible. The solid line in Fig. 3 is a fit to the logarithmic divergence predicted by both mean-field and renormalization-group theory for short-range critical wetting^{1–4}. Attempts to fit the divergence to a power-law form result in a power-law exponent of 0 (with an upper limit of 0.3), confirming the logarithmic divergence.

Contact-angle measurements close to the wetting transitions (Fig. 4) confirm the unexpected results for the layer thickness. The order of the wetting transition is determined by the discontinuity in the derivatives of the surface free energy, which is proportional to $\cos\theta$ (ref. 4). As $T \rightarrow T_w$, θ goes to zero as follows: $1 - \cos\theta \propto (T_w - T)^{2-\alpha_s}$. The surface specific-heat exponent, α_s , thus determines the order of the wetting transition. For $\alpha_s = 1$, the

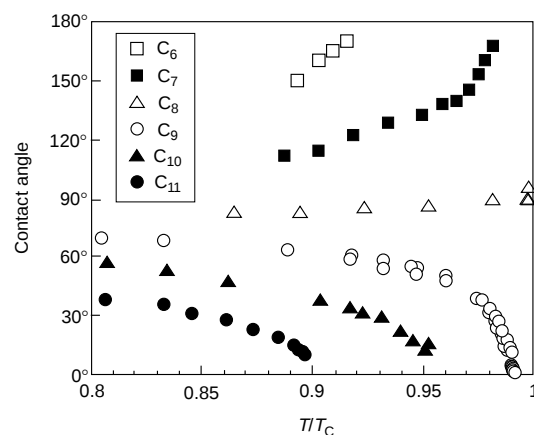


Figure 2 Contact angles of methanol on the *n*-alkanes. The measurements are taken directly from photographs similar to those shown in Fig. 1.

first derivative of $\cos\theta$ (and therefore the first derivative of the surface free energy) is discontinuous: the wetting transition is first order. For $\alpha_s < 1$, the first derivative of $\cos\theta$ is continuous: the transition is critical, implying a continuous divergence of the film thickness.

For undecane, the contact-angle data yield $\alpha_s = 1.0 \pm 0.2$, agreeing with the first-order character of the transition. In contrast, for nonane we find $\alpha_s = -0.6 \pm 0.6$, indicating a critical wetting transition. Mean-field theory predicts $\alpha_s = 0$ for short-range critical wetting^{1–4}. Renormalization-group calculations predict non-universal values for the exponent α_s , for our case $\alpha_s \approx -4.7$, in striking disagreement with the data. The renormalization-group predictions depend on the value of the fluctuation parameter, given by $\omega = k_B T_w / 4\pi\sigma\xi^2$ where k_B is Boltzmann's constant, σ is the surface tension of the alkane–methanol interface, and ξ is the bulk correlation length^{1–4}. For $T_w \rightarrow T_c$, as in our experiments, ω approaches a universal constant, $\omega_c \approx 0.7$ (ref. 13), which leads to the predicted α_s . Monte Carlo simulations^{5,6} of the transition indicate $\alpha_s = -0.8 \pm 0.4$, also in disagreement with the renormalization-group predictions, but agreeing with our experimental results.

Potential explanations for the apparent failure of the renormalization-group theory are that the transition is driven first order by fluctuations^{11,12}, that the experiments were not able to extend sufficiently close to the transition¹⁴, or that there is an error in the renormalization-group analysis. The present experiment indicates that it would be very difficult, in practice, to ascertain the cause of the failure. Given the slowness of the observed logarithmic divergence, a film thickness only twice as large as that observed here would occur just 10^{-3} K below T_w , and at a contact angle of 0.002° , making impossible observation significantly closer to the transition.

The relative importance of the mean-field and fluctuation contributions to the free energy follows from the disjoining pressure, the interaction force per unit area between the two interfaces:

$$\Pi(l) = -|A|/\xi \exp[-l/\xi] + |B|/\xi \exp[-2l/\xi] + |C|/l \exp[-l^2/\lambda^2] + |W|/l^3$$

where l is the thickness of the film; A , B and C are amplitudes; and W is a Hamaker constant, giving the strength of the long-range van der Waals interaction. The disjoining pressure is (minus) the derivative of the surface free energy, and is thus zero in equilibrium. If only the first two terms, which follow from mean-field theory (neglecting the fluctuations), are considered, a logarithmic divergence of the film thickness^{1–4} $l \propto \ln|B/A|$ is found at the critical wetting transition, for which the attractive term vanishes according to $A(T) \approx (T_w - T)$.

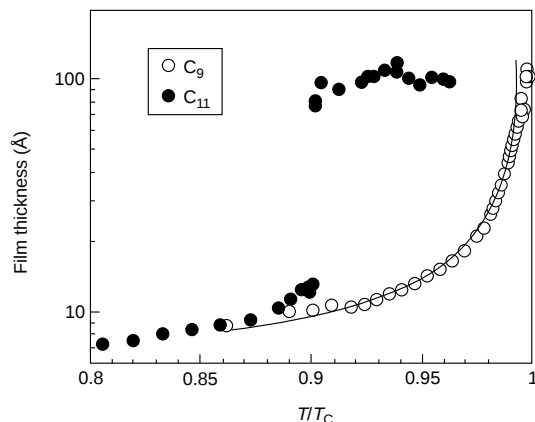


Figure 3 Equilibrium wetting film thickness of methanol. Measurements at the liquid–vapour interface of nonane (C_9) and undecane (C_{11}) are determined from ellipsometry experiments. Details of the techniques and apparatus have been described previously¹⁸. At each point, the temperature of the sample was held constant until the film thickness reached a stable and constant value. Typical equilibrium times ranged from several hours (for thin films) to several days (for the thickest films). The solid line is a fit to a logarithmic divergence of the thickness as T_w is approached, as predicted by mean-field theory. Renormalization-group theory predicts corrections on this divergence that are so small that they cannot be observed within the accuracy of the experiment.

However, because the dimensionality of the system is equal to the upper critical dimension, it is not justifiable to neglect the fluctuations^{1–4}, which are represented by the third term of $\Pi(l)$. The importance of the fluctuations is determined by the fluctuation parameter ω . For methanol/nonane ($\omega \approx 0.7$) renormalization-group analysis^{1,2} shows that the dominant repulsion between the two interfaces is no longer given by the second term in $\Pi(l)$, but rather by an entropic repulsion; this results from the confinement of the capillary-wave fluctuations of the critical (methanol–alkane) interface due to the presence of the other (methanol–vapour) interface^{1–4}.

The fluctuation term is readily calculated from capillary-wave theory^{4,15} using a short-wavelength cut-off equal to the correlation length¹⁶: $q_{\max} = \pi/\xi$. The film-thickness dependence enters via the long-wavelength cut-off, provided naturally by the presence of the other interface⁴. This leads to the third term of $\Pi(l)$, with an amplitude $C = 2\pi\sigma\omega$, and a decay length for the fluctuation repulsion given by $\lambda = \xi(6\omega)^{1/2}$. Inclusion of this term still gives critical wetting for $A(T)$ going to zero and thus does not alter the wetting temperature. In addition, renormalization-group calculations again predict a logarithmic divergence of the film thickness with the addition of a very small, $\ln(\ln|A|)$, correction^{1–4}.

It is usually assumed that short-range critical wetting cannot occur in systems with long-range (van der Waals) forces¹⁷ because these cause the transition to be of first order. It is therefore surprising to find such a transition in a real system. Inclusion of the repulsive van der Waals contribution (the fourth term in $\Pi(l)$), using estimates of the surface tension and Hamaker constant at T_w ($\sigma \approx 0.04 \text{ mN m}^{-1}$, $W \approx 10^{-2} k_B T$), reveals that the van der Waals term is always two orders of magnitude smaller than the fluctuation repulsion for nonane–methanol in the thickness range of our experiment. Thus, as foreseen by Brézin *et al.*^{1,2}, short-range critical wetting becomes possible close enough to the critical point because the long-range forces may be neglected. On the other hand, for undecane–methanol, the wetting transition occurs much farther from the critical point. This implies that the fluctuation repulsion is smaller, and that the Hamaker constant larger, so that the usual first-order transition is observed.

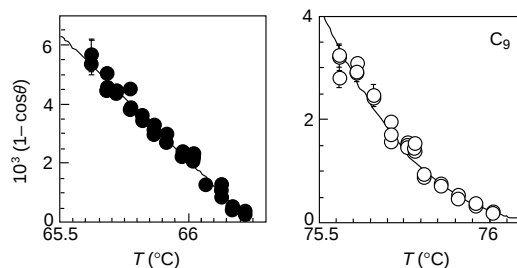


Figure 4 Contact angles of methanol on nonane (C_9) and undecane (C_{11}). Data are obtained very close to the wetting transition temperature. Measurements were made using an interferometric technique: the droplets were illuminated with monochromatic light and imaged from above with a microscope; the contact angles were computed from the observed interference pattern. Equilibration times were similar to those for Fig. 3. The solid lines are fits to the data of the form described in the text. Stated errors indicate 95% confidence intervals from the least-square fits to the data plus estimates of systematic errors.

For the observed continuous transition, the disagreement with renormalization-group theory poses a potentially serious problem, as the concept of ‘universality’, which allows for a unified description of phase transitions, relies on the correct prediction of the critical behaviour by this theory. For bulk transitions, at the upper critical dimension mean-field theory gives correct results, with small logarithmic corrections. For the surface transition presented here, mean-field theory should, according to the renormalization-group predictions, not work. We do nevertheless find mean-field behaviour. The apparent breakdown of the renormalization-group results could be due to either a false starting point for the theory, or to a critical domain that is very small. The domain of validity of mean-field theory is then the whole temperature domain that is accessible experimentally; further experiments are needed, using systems in which the fluctuation parameter can be varied, to see whether the mean-field picture eventually breaks down, and the results of the renormalization-group analysis become correct. □

Received 16 February; accepted 21 June 1999.

- Brézin, E., Halperin, B. I. & Leibler, S. Critical wetting in three dimensions. *Phys. Rev. Lett.* **50**, 1387–1390 (1983).
- Brézin, E., Halperin, B. I. & Leibler, S. Critical wetting: the domain of validity of mean field theory. *J. Phys. (France)* **44**, 775–783 (1983).
- Fisher, D. S. & Huse, D. A. Wetting transitions: a functional renormalization-group approach. *Phys. Rev. B* **32**, 247–256 (1985).
- Schick, M. in *Les Houches Summer School Session XLVIII, Liquids at Interfaces* (eds Charvolin, J., Joanny, J. F. & Zinn-Justin, J.) 415–497 (Elsevier, Amsterdam, 1990).
- Binder, K., Landau, D. P. & Kroll, D. M. Critical wetting with short-range forces: is mean-field theory valid? *Phys. Rev. Lett.* **56**, 2272–2275 (1986).
- Parry, A. O., Evans, R. & Binder, K. Critical amplitude ratios for critical wetting in three dimensions: observation of nonclassical behavior in the Ising model. *Phys. Rev. B* **43**, 11535–11538 (1991).
- Fisher, M. E. & Wen, H. Interfacial stiffness and the wetting parameter: the simple cubic Ising model. *Phys. Rev. Lett.* **68**, 3654 (1992).
- Binder, K., Landau, D. P. & Kroll, D. M. Reply. *Phys. Rev. Lett.* **68**, 3655 (1992).
- Evans, R., Hoyle, D. C. & Parry, A. O. Length scales for wetting transitions: beyond the continuum Landau approximation for the interfacial binding potential. *Phys. Rev. A* **45**, 3823–3830 (1992).
- Fisher, M. E. & Jin, A. J. Effective potentials, constraints, and critical wetting theory. *Phys. Rev. B* **44**, 1430–1433 (1991).
- Fisher, M. E. & Jin, A. J. Is short-range “critical” wetting a first-order transition? *Phys. Rev. Lett.* **69**, 792–795 (1992).
- Boulter, C. J. When is short-range “critical” wetting not critical? *Phys. Rev. Lett.* **79**, 1897–1900 (1997).
- Chaar, H., Moldover, M. R. & Schmidt, J. W. Universal amplitude ratios and the interfacial tension near consolute points of binary liquid mixtures. *J. Chem. Phys.* **85**, 418–427 (1986).
- Halpin-Healy, T. & Brézin, E. Critical wetting in three dimensions: a Ginzburg criterion. *Phys. Rev. Lett.* **58**, 1220–1223 (1987).
- Helfrich, W. in *Les Houches Summer School Session XLVIII, Liquids at Interfaces* (eds Charvolin, J., Joanny, J. F. & Zinn-Justin, J.) 209–236 (Elsevier, Amsterdam, 1990).
- Buff, P. F., Lovett, R. A. & Stillinger, F. H. Interfacial density profile for fluids in the critical region. *Phys. Rev. Lett.* **15**, 621–623 (1965).
- Nightingale, M. P., Saam, W. F. & Schick, M. Absence of critical wetting in systems with long-range forces. *Phys. Rev. Lett.* **51**, 1275–1278 (1983).
- Bonn, D., Kellay, H. & Wegdam, G. H. Experimental observation of hysteresis in a wetting transition. *Phys. Rev. Lett.* **69**, 1975–1978 (1992).

Acknowledgements. We thank E. Brézin, M. Bonn, C. Boulter, A. Parry and J. Indekeu for discussions. This work was supported by the EC.

Correspondence and requests for materials should be addressed to D.R. (e-mail: David.Ross@physique.ens.fr).

Rapid environmental changes in southern Europe during the last glacial period

Judy R. M. Allen*, Ute Brandt†, Achim Brauer†‡, Hans-Wolfgang Hubberten§, Brian Huntley*, Jörg Keller||, Michael Kramll, Andreas Mackensen¶, Jens Mingram†, Jörg F. W. Negendank†, Norbert R. Nowaczyk†, Hedi Oberhänsli†, William A. Watts#, Sabine Wulff† & Bernd Zolitschka†

* Environmental Research Centre, University of Durham, Durham DH1 3LE, UK

† GeoForschungsZentrum Potsdam, Projektbereich 3.3, Telegrafenberg, D-14473 Potsdam, Germany

‡ Faculté des Sciences et Techniques de Saint-Jérôme, Laboratoire de Botanique Historique et Palynologie, F-13379 Marseille Cedex 20, France

§ Alfred-Wegener-Institut für Polar- und Meeresforschung, Telegrafenberg, D-14473 Potsdam, Germany

|| Institut für Mineralogie, Petrologie und Geochemie, Albertstrasse 23b, D-79104 Freiburg im Breisgau, Germany

¶ Alfred-Wegener-Institut für Polar- und Meeresforschung, Columbus-Zentrum, D-27515 Bremerhaven, Germany

Department of Botany, Trinity College, Dublin 2, Ireland

Oxygen-isotope records from Greenland ice cores^{1,2} indicate numerous rapid climate fluctuations during the last glacial period. North Atlantic marine sediment cores show comparable variability in sea surface temperature and the deposition of ice-rafted debris^{3–5}. In contrast, very few continental records of this time period provide the temporal resolution and environmental sensitivity necessary to reveal the extent and effects of these environmental fluctuations on the continents. Here we present high-resolution geochemical, physical and pollen data from lake sediments in Italy and from a Mediterranean sediment core, linked by a common tephrochronology. Our lacustrine sequence extends to the past 102,000 years. Many of its features correlate well with the Greenland ice-core records, demonstrating that the closely coupled ocean–atmosphere system of the Northern Hemisphere during the last glacial⁴ extended its influence at least as far as the central Mediterranean region. Numerous vegetation changes were rapid, frequently occurring in less than 200 years, showing that the terrestrial biosphere participated fully in last-glacial climate variability. Earlier than 65,000 years ago, our record shows more climate fluctuations than are apparent in the Greenland ice cores. Together, the multi-proxy data from the continental and marine records reveal differences in the seasonal character of climate during successive interstadials, and provide a step towards determining the underlying mechanisms of the centennial–millennial-scale variability.

A series of four sediment cores (B, D, J and L) obtained from Lago Grande di Monticchio (40° 56' N, 15° 35' E, 656 m above sea level), a maar lake in Basilicata, southern Italy, extends to a depth of 72.5 m. Sedimentation rates, estimated from annually laminated sections of a composite of these cores, provide a chronology^{6,7} that gives a date of 101.7 kyr ago for the base of the record (Fig. 1). This calendar-year chronology, based solely upon Monticchio sedimentation rates, is independent of palynostratigraphic (that is, pollen-based), marine $\delta^{18}\text{O}$ event or ice-core interstadial correlations. It is complemented by a tephrochronology and a series of radioisotopic ages. A total of 340 tephra layers has been recorded, the more distinctive of which have been labelled as 'marker tephtras' (MT-1 to MT-11) (Table 1; see also Supplementary Information). These marker tephtras, originating from the Campanian igneous province, along with unambiguous sediment features, provide the basis for linking the series of cores. Several of the thicker tephtras have been correlated

with radioisotopically-dated tephtras elsewhere in Italy or in sediment cores from the Mediterranean Sea^{8,9} (Table 1). $^{40}\text{Ar}/^{39}\text{Ar}$ age determinations also have been made upon three of the tephtras from Monticchio¹⁰. ^{14}C determinations are problematic due to volcano-genic CO_2 emissions into the lake; potentially more reliable ^{14}C age determinations have, however, been obtained using terrestrial macrofossils¹¹. Almost all of these tephtra and radioisotopic dates concur, to within 1σ uncertainty, with the sedimentation-rate chronology^{6,7}. In the case of the $^{40}\text{Ar}/^{39}\text{Ar}$ age of 105 ± 2 kyr, determined for the X-5 tephtra in core M25/4-12 from the Ionian basin¹² (Fig. 1) and assigned to the MT-10 tephtra on the basis of mineralogical and geochemical evidence, the 2σ uncertainty range encompasses the Monticchio age.

The pollen spectra were used to reconstruct the main vegetation units (biomes¹³); these are indicated alongside the simplified pollen diagram (Fig. 2). The principal features of this record are: (1) the Holocene (11.2 kyr ago to the present), characterized by warm mixed and temperate deciduous forest biomes; (2) the Last Glacial Maximum (pollen assemblage zone (PAZ) 4; 25.9–14.3 kyr ago), other last-glacial stadials (especially PAZs 6, 8, 10, 12, 14 and 16) and PAZ 18 (86.8–84.3 kyr ago), characterized by steppe biomes; (3) last-glacial interstadials, characterized principally by the wooded steppe biome, in which a diverse range of temperate trees was present, with temperate deciduous and cool mixed forest biomes occurring only rarely; and (4) PAZs 17 and 19 (84.2–72.0 and 101.7–86.8 kyr ago), characterized by predominance of the temperate deciduous forest biome. A striking general feature of the pollen record is the rapidity with which forest or wooded steppe biomes replace steppe biomes and vice versa; the mean interval for absolute changes of $>20\%$ in total pollen of woody taxa is 142 yr (standard error (s.e.) = 21, $n = 46$), with decreases typically occur-

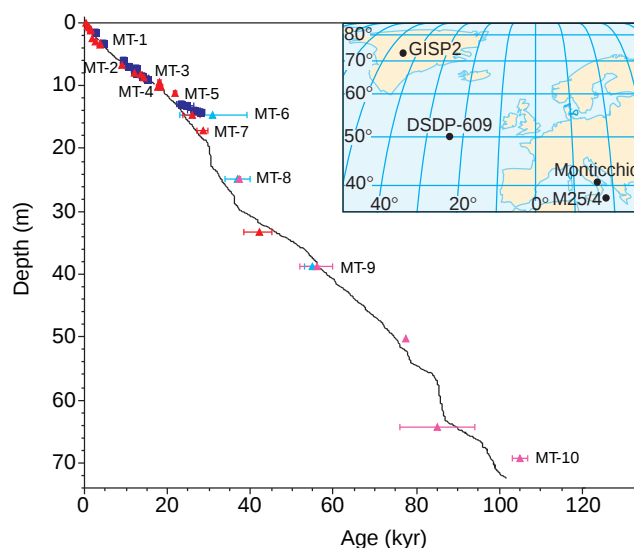


Figure 1 Site chronology. Main figure, age–depth relationship based on the sedimentation-rate chronology (black line). This chronology is derived from 1,120 sedimentation-rate estimates for cores B, D and J, obtained from counts and thickness measurements of annual laminations present in most sediment segments (B and D above 75,610 yr, and J below this). These measurements were made on thin sections of sediment, examination of which also enabled the annual nature of the laminations to be established⁶ (data are available as Supplementary Information, and will also be made available at the World Data Center A for Paleoclimatology). Also shown are radioisotopic dates obtained from Monticchio^{10,11}—calibrated ^{14}C sediment ages (dark blue squares) and $^{40}\text{Ar}/^{39}\text{Ar}$ tephtra ages (pale blue triangles)—and ages for tephtras matched geochemically and mineralogically to pyroclastics dated in other terrestrial or marine sediment records (Table 1)— $^{40}\text{Ar}/^{39}\text{Ar}$ ages of tephtras (pink triangles)^{12,20,23} and calibrated ^{14}C ages of incorporated or underlying organic materials (red triangles)^{16,18,19,24–26}. Inset, locations of sites studied and others mentioned in the text.

Table 1 Origin and ages of marker tephra layers

Tephra	Origin of tephra*	Mediterranean marine equivalent ^a	Monticchio age†	Reported age‡
MT-1	Vesuvius	Z-1	4,020	3,760 (70) ¹⁶
MT-2	Vesuvius	Mercato	9,680	8,890 (120) ¹⁶
MT-3	Phlegrean Fields	Pomici Principali	12,180	12,170 (400) ¹⁷
MT-4	Vesuvius	Greenish	17,560	17,830 (300) ¹⁶
MT-5	Vesuvius	Pomici di Base	19,280	21,740 (390) ¹⁶
MT-6	Campanian province	Sarno	23,930	25,820 (2,080) ¹⁸
MT-7	Vesuvius	Codola	26,790	28,500 (1,280) ¹⁹
MT-8	Phlegrean Fields	Y-5	32,970	37,100 (400) ²⁰
MT-9	Ischia	Y-7	56,250	56,000 (4,000) ¹²
MT-10	Campanian province	X-5	97,770	105,000 (2,000) ¹²
MT-11	Campanian province	Unknown	98,750	Undated

* Narcisi²¹ previously correlated tephra MT-1 to MT-9 with terrestrial volcanic deposits in Italy, and in some cases with marine tephra records.

† Calendar-year chronology^{6,7} based on annual laminations and interpolated sedimentation rates (Fig. 1).

‡ MT-1 to MT-7: ¹⁴C ages in years calibrated following refs 21 and 22; MT-8 to MT-10: multiple single-crystal and multigrain aliquot ⁴⁰Ar/³⁹Ar measurements (in years; s.d. in parentheses).

ring more quickly than increases.

The main features of the stratigraphic records of physical properties, dry density (Fig. 3a) and magnetic susceptibility (Fig. 3b), generally parallel those for relative abundance of pollen of woody taxa (Fig. 3e). Abundance of biogenic silica (Fig. 3d), derived principally from diatoms, exhibits the same trend. Estimated organic carbon content (Fig. 3c) is high during the Holocene and low throughout the remainder of the record, with only slight increases during the earlier forest periods. Thus intervals of high

abundance of pollen of woody taxa are typified by sediments of low dry density and magnetic susceptibility, and high biogenic silica content; intervals dominated by herbaceous pollen taxa have the opposite sediment characteristics. Together these characteristics indicate a more stable catchment, with less erosion of mineral materials, and a more productive lake, during times of forest biomes. The intervals characterized by steppe biomes, in contrast, have less stable soils and less productivity within the lake. Given the present association of steppes with climates characterized by moderate aridity and seasonal droughts, enhanced soil erosion might result from spatially and/or temporally discontinuous vegetation cover. Reduced productivity within the lake reflects both colder conditions and reduced nutrient availability, the latter a consequence of decreased chemical weathering related to lower temperatures and precipitation.

Although the general pattern of environmental changes can be inferred qualitatively from changes in the reconstructed biomes (Fig. 2), quantitative palaeoclimate reconstructions made from pollen data¹⁴ (Fig. 3f–h) provide greater insight into the climate changes. These reconstructions indicate that moisture availability (Fig. 3h) was reduced during the cold/warm steppe periods compared to the forest periods, with the wooded steppe periods intermediate. Although consistently close to its modern value during forested periods, the mean temperature of the coldest month (Fig. 3f) varied markedly during the steppe periods. Whereas extreme low values were >20 °C less than modern, during most of the last glacial values were intermediate and ~12 °C less than today.

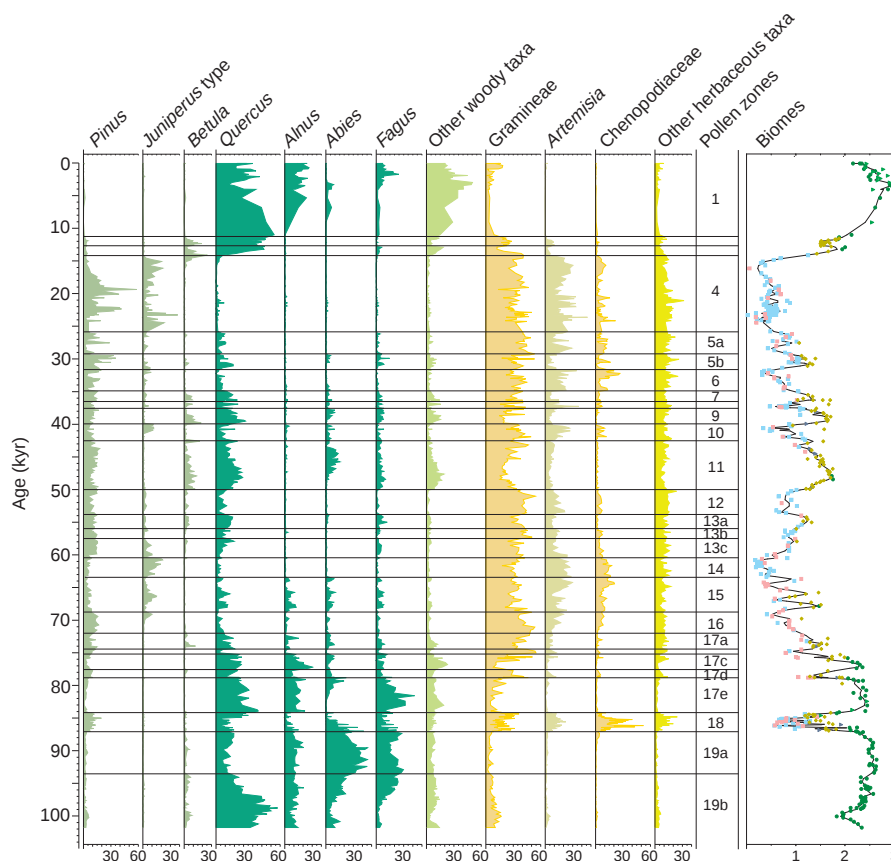


Figure 2 Vegetation change at Monticchio. On the left is a diagram showing the contribution (in %) of selected pollen taxa to the total pollen from Lago Grande di Monticchio cores D and J. A mean of 399 (standard deviation, 130) grains of terrestrial pollen was counted; percentages were calculated with respect to this sum. Pollen zones are based on stratigraphically constrained incremental sum of squares analysis. Right panel: biomes were calculated using all terrestrial pollen taxa^{13,27} and are plotted (data points) against sample scores on the first detrended correspondence analysis axis of the data set (eigenvalue 0.39). Also shown

(black line) are the first-axis scores smoothed using a 'lowess' procedure²⁸ with bandwidth 0.01. Three forest and three steppe biomes occur: warm mixed forest (green triangles), temperate deciduous forest (green circles), cool mixed forest (blue triangles), wooded steppe (brown diamonds), warm steppe (pink squares) and cold steppe (pale blue squares). (Data for this figure are available as Supplementary Information, and will also be made available at the World Data Center A for Paleoclimatology.)

The annual temperature sum (Fig. 3g) varied more than the other two climate variables within the forested periods, notably falling during PAZ 19a compared to the preceding sub-zone. The very rapid vegetation changes are reflected by rapid changes in reconstructed climate; the mean temperature of the coldest month in particular exhibits absolute changes of $>10^{\circ}\text{C}$ within a mean interval of 153 yr (s.e. = 54.2, $n = 8$). Numerous rapid climate fluctuations of relatively large magnitude occurred during the last glacial. Of particular note are two features earlier than 72 kyr ago: (1) the numerous very rapid large-magnitude fluctuations during PAZ 18; and (2) the series of large-magnitude fluctuations during PAZ 17.

The high-resolution oxygen-isotope stratigraphy from marine core M25/4-11 (Fig. 1), dated by tephrochronology of this and the adjacent core M25/4-12, enables examination of correlations between Monticchio and the marine record (Fig. 3i). It is apparent that sapropels in Mediterranean Sea sediments (Fig. 3i) typically correspond to periods of relatively high annual temperature sum and high lacustrine productivity (Fig. 3c, d, g) that in turn correspond to maxima in July insolation (Fig. 3m). In North Atlantic cores (for example, DSDP-609, Fig. 1) the relative abundance of planktonic foraminiferan *Neogloboquadrina pachyderma* (s.), negatively correlated with sea surface temperature, fluctuates rapidly and repeatedly during the last glacial (Fig. 3k). Although

particularly large increases are associated with peaks in ice-rafted debris (Fig. 3j) known as Heinrich events^{3,5}, these events are marked by minima in absolute abundance of *N. pachyderma* (s.). Smaller-amplitude variations in its relative abundance, and hence in sea surface temperature, have been correlated with Dansgaard-Oeschger cycles seen in the $\delta^{18}\text{O}$ record from Greenland ice cores (Fig. 3l)^{1,2}, leading to the inference that the ocean-atmosphere system was closely coupled during the last glacial⁴. The records of *N. pachyderma* (s.) from DSDP-609 (Fig. 3k) and of $\delta^{18}\text{O}$ from GISP2 (Fig. 3l) can be correlated with the Monticchio record (Fig. 3a–h), demonstrating a link between the Atlantic climate system and the central Mediterranean region. Correlations become more difficult before ~65 kyr ago, perhaps reflecting dating problems in the DSDP-609 and/or GISP2 records as a result of the age modelling approaches applied^{1,4}.

Although correlates for the interstadials recognised in GISP2 (Fig. 3l) can be identified in the Monticchio record, before 65 kyr ago the latter shows far more detail than the ice-core data with evidence of additional fluctuations. In addition, the availability of multi-proxy data from Monticchio reveals the complexity of the environmental changes. For example, interstadials 19 and 20 have similar characteristics in the GISP2 $\delta^{18}\text{O}$ record but differ markedly at Monticchio. During interstadial 19 (correlated with PAZ 15), biogenic silica values increase only slightly, accompanied by a small

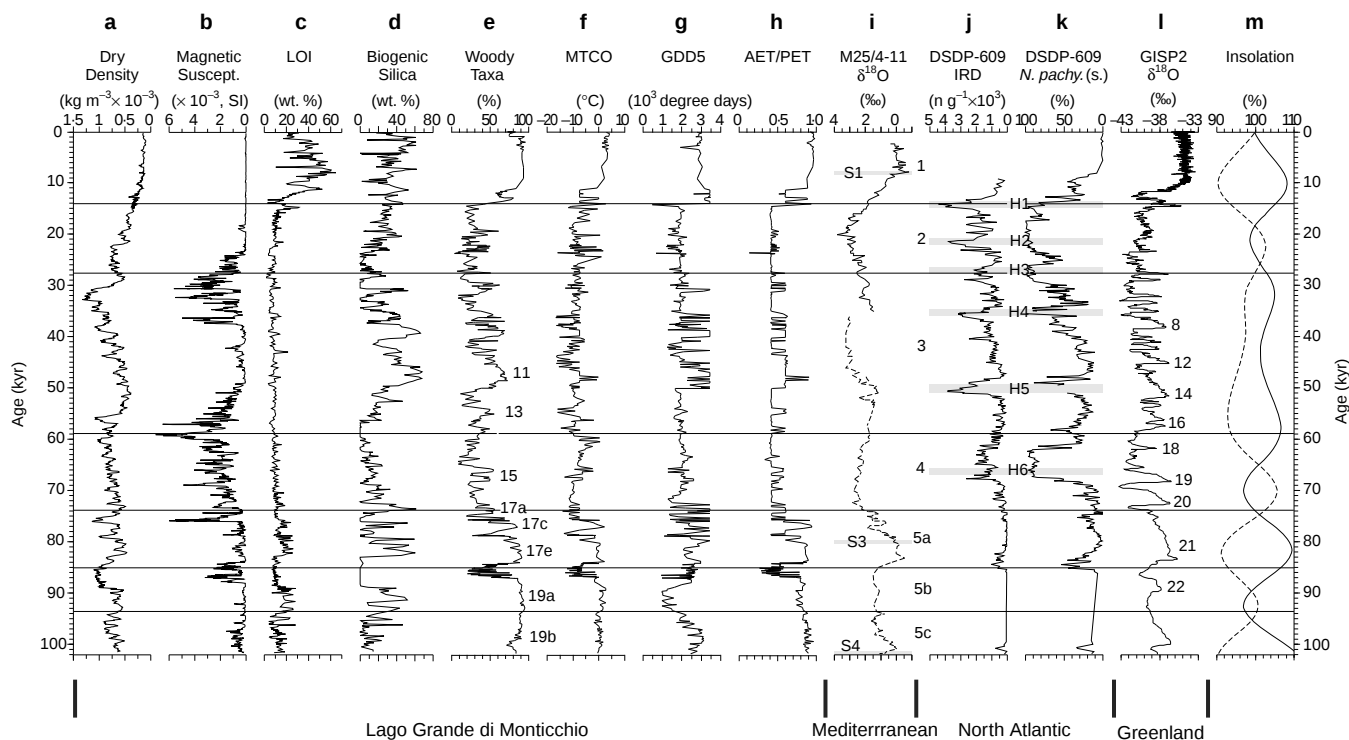


Figure 3 Palaeoenvironmental records of the last glacial period. Sediment parameters and palaeoclimate reconstructions from Monticchio are compared with marine and ice-core records (all curves plotted on their own calendar-year timescale). **a**, Dry density (smoothed). **b**, Magnetic susceptibility (smoothed). **c**, Organic carbon content estimated by loss in weight upon ignition at 550°C (LOI). **d**, Biogenic silica estimated using a normative model based on X-ray fluorescence major-element data²⁹; negative modelled values have been set to zero (76 kyr to present, after ref. 29). **e**, Relative abundance of pollen of woody taxa and numbers of pollen zones discussed in the text. **f**, Mean temperature of the coldest month (MTCO). **g**, annual temperature sum above 5°C (GDD5) and **h**, estimate of the ratio of actual to potential evapotranspiration (AET/PET) reconstructed from the pollen data using pollen-climate response surfaces¹⁴. **i**, $\delta^{18}\text{O}$ ‰ PDB record for marine sediment core M25/4-11 from the Ionian basin of the Mediterranean Sea measured on planktonic foraminiferans (*Globigerinoides ruber*, solid line; *Globigerina bulloides*, dashed line). Shaded bars indicate

sapropels (S1, S3 and S4). **j**, Abundance of lithic grains $>150\mu\text{m}$ (n g^{-1}) (IRD) and **k**, relative abundance of *Neogloboquadrina pachyderma* (s.) for North Atlantic core DSDP-609^{3,5}. Shaded bars indicate periods of increased abundance of lithic grains related to Heinrich events (H1–H6)^{3,5}. **l**, Greenland ice-core (GISP2) $\delta^{18}\text{O}$ ‰ SMOW record¹ with selected interstadials marked. **m**, January (dashed line) and July (solid line) insolation at 40°N expressed as percentages of present values (calculated at 250-yr intervals using the program developed by Berger³⁰). Horizontal dashed lines indicate MIS stage and sub-stage boundaries (ages after ref. 28) and the MIS stage numbers are shown between curves in **i** and **j**. Data used to plot curves in **j**, **k** and **l** were obtained from World Data Center A (WDC-A), operated by the National Geophysical Data Center at Boulder; data for the remaining curves, with the exception of **m**, are available as Supplementary Information, and will also be made available at the World Data Center A for Paleoclimatology. Note reversed x-axis scales for panels **a**, **b**, **i**, **j** and **k**.

increase in pollen of woody taxa; during interstadial 20 (correlated with PAZ 17a), biogenic silica values show a strong peak (Fig. 3d) while pollen of woody taxa is similar in abundance to that in interstadial 19 (Fig. 3e). Environmentally, the two interstadials were apparently equally moist, but whereas interstadial 19 was warmer in winter than interstadial 20, the latter had a higher annual temperature sum than the former; such differences almost certainly reflect differing atmospheric circulation patterns during the two events. Characteristics of the additional environmental fluctuations can also be inferred. Thus during interstadial 21, corresponding to PAZs 17e–c, the decrease in abundance of pollen of woody taxa during PAZ 17d and the corresponding period of lower lake productivity indicate conditions comparable to those during later stadials. This event is recorded as an increase in $\delta^{18}\text{O}$ of planktonic foraminiferans in core M25/4-11 (Fig. 3i), but has no correlate in the GISP2 record. PAZ 18, a period of rapid environmental fluctuations at Monticchio, is represented in the GISP2 record by a decrease in $\delta^{18}\text{O}$ values between interstadials 22 and 21 and in DSDP-609 by increased, and to some extent fluctuating, relative abundance of *N. pachyderma* (s.) (Fig. 3k).

Our record from Monticchio demonstrates the capability of late Quaternary lake sediments to provide sensitive, high-resolution records of rapid (centennial–millennial) environmental fluctuations comparable to those obtained from ice cores. It also reveals that the biosphere was a full participant in these rapid fluctuations, contrary to widely held views that vegetation is unable to change with such rapidity. The opportunity to develop an independent calendar-year chronology (Fig. 1) allows comparison to precisely dated records from other realms without relying on correlating their principal features ('wiggles matching'); this allows us to quantify, for example, the much shorter duration in terrestrial records of the stadial event correlated with marine oxygen-isotope substage 5b¹⁵. Lake sediments also have the advantage of recording many proxies of past environments, allowing seasonal climate characteristics to be reconstructed. The Monticchio record demonstrates that the closely coupled Northern Hemisphere ocean–atmosphere system of the last glacial period⁴ extended its influence beyond the North Atlantic and Greenland, at least as far as the central Mediterranean region. In addition, the multiple proxies reveal differences in the character of the climate during successive interstadials, as well as revealing additional climate fluctuations before 65 kyr ago not evident in records from other archives, most probably because of limitations of the ice cores in particular. Although, given predominant mid-latitude atmospheric circulation patterns, the linkage of the Mediterranean and North Atlantic regions during the last glacial should come as no surprise, the new information about the varying character and expression of fluctuations in the Mediterranean region provides a first step towards discriminating between alternative mechanisms that might have generated millennial-scale variability during the last glacial period. □

Received 17 February; accepted 9 July 1999.

- Groote, P. M., Stuiver, M. White, J. W. C., Johnsen, S. & Jouzel, J. Comparison of oxygen isotope records from the GISP2 and GRIP Greenland ice cores. *Nature* **366**, 552–554 (1993).
- Dansgaard, W. *et al.* Evidence for general instability of past climate from a 250-kyr ice-core record. *Nature* **364**, 218–220 (1993).
- Bond, G. *et al.* Evidence for massive discharges of icebergs into the North Atlantic ocean during the last glacial period. *Nature* **360**, 245–249 (1992).
- Bond, G. *et al.* Correlations between climate records from North Atlantic sediments and Greenland ice. *Nature* **365**, 143–147 (1993).
- Bond, G. C. & Lotti, R. Iceberg discharges into the North Atlantic on millennial time scales during the last glaciation. *Science* **276**, 1005–1010 (1995).
- Zolitschka, B. & Negendank, J. F. W. Sedimentology, dating and palaeoclimatic interpretation of a 76.3 ka record from Lago Grande di Monticchio, southern Italy. *Quat. Sci. Rev.* **15**, 101–112 (1996).
- Zolitschka, B. *Palaoklimatische Bedeutung laminierter Sedimente—Holzmaar (Eifel, Deutschland), Lake C2 (Nordwest-Territorien, Kanada) und Lago Grande di Monticchio (Basilikata, Italien)* (Gebauer Bornträger, Berlin, 1998).
- Narcisi, B. Tephrochronology of a Late Quaternary lacustrine record from the Monticchio maar (Vulture volcano, southern Italy). *Quat. Sci. Rev.* **15**, 155–166 (1996).
- Keller, J., Ryan, W. B. F., Ninkovich, D. & Altherr, R. Explosive volcanic activity in the Mediterranean over the past 200,000 yr as recorded in deep-sea sediments. *Geol. Soc. Am. Bull.* **89**, 591–604 (1978).
- Watts, W. A., Allen, J. R. M. & Huntley, B. Vegetation history and palaeoclimate of the last glacial period at Lago Grande di Monticchio, southern Italy. *Quat. Sci. Rev.* **15**, 133–154 (1996).

- Hajdas, B., Bonani, G., Zolitschka, B., Brauer, A. & Negendank, J. ¹⁴C ages of terrestrial macrofossils from Lago Grande di Monticchio (Italy). *Radiocarbon* **40**, 803–807 (1998).
- Kraml, M. *Laser-⁴⁰Ar/³⁹Ar-Datierungen an distalen marinen Tephren des jung-quartären mediterranen Vulkanismus (Ionisches Meer, METEOR-Fahrt 25/4)*. Thesis, Univ. Freiburg (1997).
- Prentice, I. C., Guiot, J., Huntley, B., Jolly, D. & Cheddadi, R. Reconstructing biomes from palaeoecological data: a general method and its application to European pollen data at 0 and 6 ka. *Clim. Dyn.* **12**, 185–194 (1996).
- Huntley, B. The use of climate response surfaces to reconstruct palaeoclimate from Quaternary pollen and plant macrofossil data. *Phil. Trans. R. Soc. Lond.* **341**, 215–223 (1993).
- Kukla, G., McManus, J. F., Rousseau, D.-D. & Chuine, I. How long and how stable was the last interglacial? *Quat. Sci. Rev.* **16**, 605–612 (1997).
- Andronico, D. *et al.* Geological map of Somma-Vesuvius volcano. *Periodico Mineral.* **64**, 77–88 (1995).
- Di Vito, M. A. *et al.* Volcanic and deformation history of the Campi Flegrei caldera in the past 12 ka. *J. Volcanol. Geotherm. Res.* (in the press).
- Capaldi, G., Civetta, L. & Gillot, P. Y. Geochronology of Plio-Pleistocene volcanic rocks from southern Italy. *Rendiconti Soc. Ital. Mineral. Petrol.* **40**, 25–44 (1985).
- Alessio, M. *et al.* University of Rome carbon-14 dates XII. *Radiocarbon* **16**, 358–367 (1974).
- Deino, A. L., Southon, J., Terras, F., Campagnola, L. & Orsi, G. ¹⁴C and ⁴⁰Ar/³⁹Ar dating of the Campanian Ignimbrite, Phlegrean Fields, Italy. *8th International Conference on Geochronology, Cosmochronology and Isotope Geology* (eds Lamphere, M. A., Dalrymple, G. B. & Turrin, B. D.) **1107**, 77 (United States Geological Survey Circular, 1994).
- Stuiver, M. *et al.* INTCAL98 radiocarbon age calibration, 24,000–0 cal BP. *Radiocarbon* **40**, 1041–1083 (1998).
- Laj, C., Mazaud, A. & Duplessy, J.-C. Geomagnetic intensity and ¹⁴C abundance in the atmosphere and ocean during the past 50 kyr. *Geophys. Res. Lett.* **23**, 2045–2048 (1996).
- Fornaseri, M. Geochronology of volcanic rocks from Latium (Italy). *Rendiconti Soc. Ital. Mineral. Petrol.* **40**, 73–106 (1985).
- Alessio, M. *et al.* University of Rome carbon-14 dates IX. *Radiocarbon* **13**, 395–411 (1971).
- Orsi, G. *et al.* ¹⁴C geochronological constraints for the volcanic history of the island of Ischia (Italy) over the last 5000 years. *J. Volcanol. Geotherm. Res.* **71**, 249–257 (1996).
- Vezzoli, L., Coltelli, M. & Del Carlo, P. Tephrochronology: from source area studies to stratigraphic and palaeoenvironmental perspectives, an example of Etna volcano (Italy). *International Union for Quaternary Research Commission on Tephrochronology Union Internationale des Sciences Pré- et Protohistoriques 31* (Inter Congress Symposium, Brives-Charensac, France, 1998).
- Tarasov, P. E. *et al.* A method to determine warm and cool steppe biomes from pollen data; application to the Mediterranean and Kazakhstan regions. *J. Quat. Sci.* **13**, 335–344 (1998).
- Whitlock, C. & Bartlein, P. J. Vegetation and climate change in northwest America during the past 125 yr. *Nature* **388**, 57–61 (1997).
- Robinson, C. Lago Grande di Monticchio, southern Italy: a long record of environmental change illustrated by sediment geochemistry. *Chem. Geol.* **118**, 235–254 (1994).
- Berger, A. Long term variations of caloric insolation resulting from the Earth's orbital elements. *Quat. Res.* **9**, 139–167 (1978).

Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from the London editorial office of Nature.

Acknowledgements. We thank A. Hendrich for drafting the figures. Initial parts of this study were supported by the EU 'Euromars' Programme. U.B., M.K., S.W. and B.Z. were supported by the Deutsche Forschungsgemeinschaft, and J.R.M.A. was supported by the UK Natural Environment Research Council. This study is a contribution to the European Lake Drilling Programme, and to the PEP III PAGES transect through Europe and Africa.

Correspondence and requests for materials should be addressed to B.H. (e-mail: Brian.Huntley@Durham.ac.uk).

Effects of ship emissions on sulphur cycling and radiative climate forcing over the ocean

Kevin Capaldo*, James J. Corbett†, Prasad Kasibhatla‡, Paul Fischbeck†§ & Spyros N. Pandis*†

Departments of * Chemical Engineering, † Engineering and Public Policy, and § Social and Decision Sciences, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, USA

‡ Nicholas School of the Environment, Duke University, Durham, North Carolina 27708, USA

The atmosphere overlying the ocean is very sensitive—physically, chemically and climatically—to air pollution. Given that clouds over the ocean are of great climatic significance, and that sulphate aerosols seem to be an important control on marine cloud formation¹, anthropogenic inputs of sulphate to the marine atmosphere could exert an important influence on climate. Recently, sulphur emissions from fossil fuel burning by international shipping have been geographically characterized², indicating that ship sulphur emissions nearly equal the natural sulphur flux from ocean to atmosphere in many areas³. Here we use a global chemical transport model to show that these ship emissions

can be a dominant contributor to atmospheric sulphur dioxide concentrations over much of the world's oceans and in several coastal regions. The ship emissions also contribute significantly to atmospheric non-seasalt sulphate concentrations over Northern Hemisphere ocean regions and parts of the Southern Pacific Ocean, and indirect radiative forcing due to ship-emitted particulate matter (sulphate plus organic material) is estimated to contribute a substantial fraction to the anthropogenic perturbation of the Earth's radiation budget. The quantification of emissions from international shipping forces a re-evaluation of our present understanding of sulphur cycling and radiative forcing over the ocean.

Ocean-going ships represent a significant global industry whose environmental impacts have not previously been evaluated quantitatively, except in limited port and regional studies⁴. In fact, the first comprehensive, geographically resolved inventory of ship emissions has only recently been completed^{2,3}; it revealed two important findings. First, over large sections of the Northern Hemisphere and in some regions of the Southern Hemisphere, anthropogenic sulphur emissions from ships are comparable to biogenic dimethyl sulphide (DMS) emissions, the dominant source of sulphur from the ocean. Second, ship emissions have the potential to affect air quality in many coastal and port regions along heavily-travelled international trade routes, where annual sulphur emissions from ships equal (or exceed) those of adjacent land-based sources.

Here we build on this work by estimating the contribution of ship emissions to ambient concentrations of SO₂, sulphate, and cloud condensation nuclei (CCN), using a data-plus-model approach. We model the atmospheric sulphur cycle with the Geophysical Fluid Dynamics Laboratory's three-dimensional global chemical transport model, which includes transport, wet and dry deposition, and gas, cloud, and aerosol phase sulphur chemistry⁵. DMS is included as a prognostic variable, to help better describe the atmospheric sulphur cycle over the ocean. Oxidation of DMS to SO₂ is represented using two reactions with the hydroxyl (OH) radical and a reaction with the nitrate radical (NO₃). Production of SO₂ from DMS is assumed to have a temperature-dependent net yield between 0.88 and 0.98 (ref. 6). Biogenic sulphur emissions, including ocean-based DMS emissions, are from Benkovitz *et al.*⁷; sulphur emissions from ships are from the SEA (Ship Emissions Assessment) inventory³; and land-based anthropogenic sulphur emissions are from the GEIA (Global Emissions Inventory Activity) data set⁴. All input data sets are seasonally resolved and gridded to a model resolution of ~265 km. GEIA estimates of shipping emissions in the northeast Atlantic Ocean were removed to avoid double counting with the global SEA sulphur inventory. To minimize sensitivity to initial species concentrations, 20 months were simulated, and the results of the last 12 months are presented here.

By comparing model results that include ship emissions with ones that do not, the relative effect of ships on ambient SO₂ and NSS-sulphate concentrations can be quantified. Model calculations indicate that ship emissions account for greater than 60% of the July SO₂ prediction for large areas of the north Atlantic and north Pacific (Fig. 1). For most of the rest of the Northern Hemisphere oceans, ship emissions are responsible for greater than 30% of the predicted SO₂. A smaller effect is seen in the Southern Hemisphere oceans, where ship emissions account for greater than 40% of the predicted SO₂ only around Australia, New Zealand, South Africa, and Argentina. NSS-sulphate predictions for July (Fig. 2) show that ship emissions account for 10% to 30% of the predicted NSS-sulphate over most of the Northern Hemisphere oceans: ships in the Arctic Ocean north of Norway and the eastern Pacific generate 30% to 50% of the predicted NSS-sulphate. In the Southern Hemisphere, ship contributions are generally less than 5%, except over large areas north and east of Australia where ships contribute between 10% and 20% of the ambient NSS-sulphate.

Seasonal variations in the model predictions depend mainly on

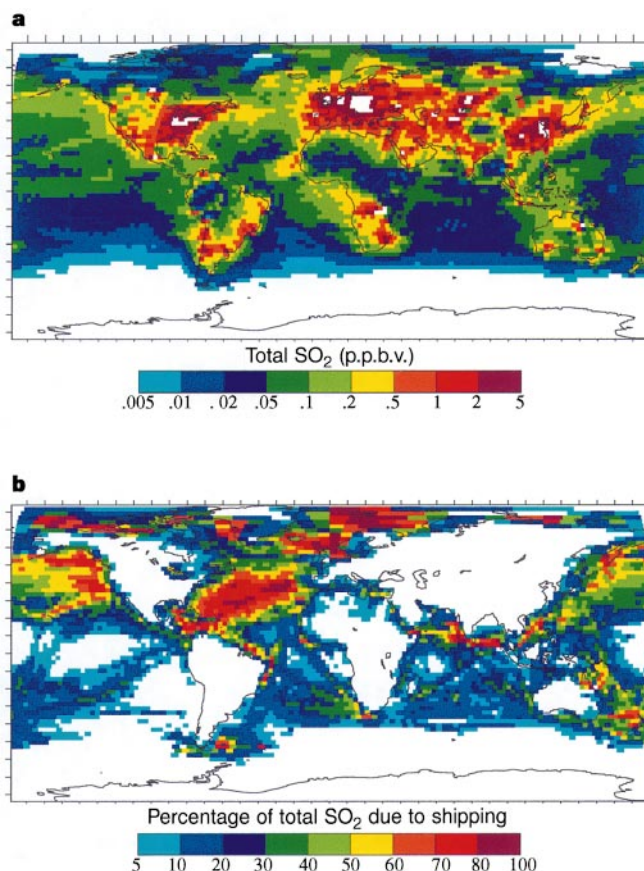


Figure 1 Contribution of ship emissions to SO₂, July. **a**, Total predicted surface concentrations (p.p.b.v.) including ship, land-based anthropogenic and global biogenic sulphur emissions. The white regions over the continents represent predictions above 5 p.p.b.v. and the white regions at the poles represent predictions below 0.005 p.p.b.v. **b**, The percentage of the total due to ship emissions. White regions represent contributions less than 5%.

changes in temperature-dependent DMS emissions, land-based energy use, and meteorology. These factors tend to dominate over the smaller seasonal differences in ship-route patterns and traffic intensity³. The decrease in DMS production during the winter increases the relative effect of ship emissions. The January contribution of ship emissions to SO₂ predictions increases to around 90% for large areas of the north Atlantic and northeast Pacific. At the same time, winter land-based emissions increase and are transported farther. This decreases the January ship contribution to SO₂ concentrations to less than 20% in areas around Japan, the east coast of the United States, and in the Atlantic north of 50° N. Because the longer-lived NSS-sulphate species is transported farther than SO₂, ships contribute less than 5% to modelled NSS-sulphate concentrations in most areas outside the tropics during winter.

Comparisons of actual SO₂ observations with model predictions highlight the importance of explicitly considering ship emissions as a factor in studies of the marine atmosphere. We compared monthly averaged, ocean-based SO₂ observations from 15 measurement campaigns^{8–15} with the corresponding model predictions for the same month and area (within the resolution of the model), excluding monthly averaged measurements greater than 300 parts per trillion (10¹²) by volume, p.p.t.v., as these appear to be influenced substantially by land-based emissions (Fig. 3). The fit between observations and model predictions for SO₂ improves when ship emissions are added (the difference in regression slopes is statistically significant ($p < 0.0001$), and the correlation coefficient with SO₂ observations improves slightly from $r = 0.51$ to $r = 0.55$). Moreover, the variability of the model predictions when ship

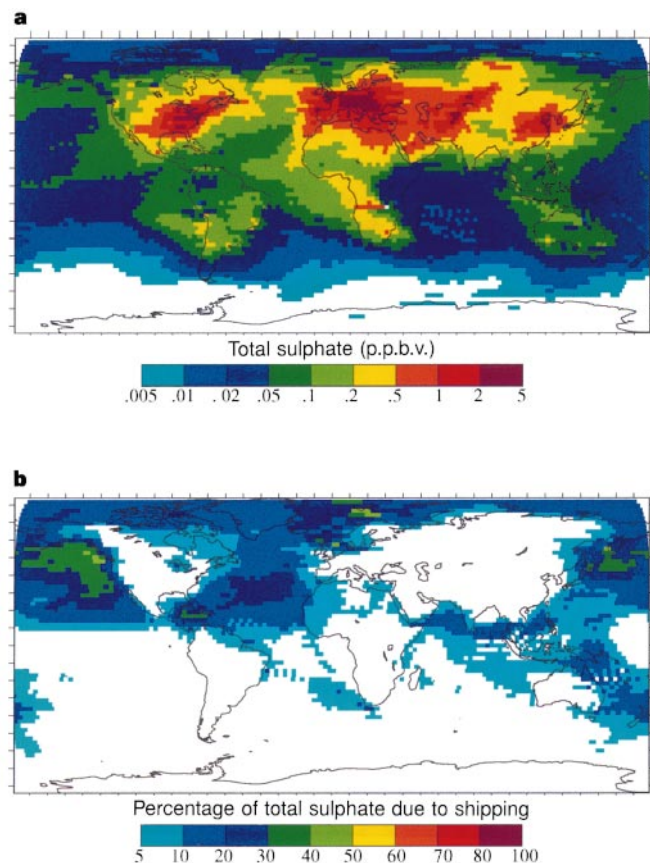


Figure 2 Contribution of ship emissions to non-seasalt sulphate, July. **a**, Total surface concentrations (p.p.b.v.) including ship, land-based anthropogenic and global biogenic sulphur emissions. **b**, The percentage of the total due to ship emissions. White regions represent regions less than 0.005 p.p.b.v. and ship contributions less than 5%.

emissions are included (standard deviation $\sigma = 36$ p.p.t.v.) is larger than (and closer to) the observed data ($\sigma = 52$ p.p.t.v.) than when ship emissions are not included ($\sigma = 27$ p.p.t.v.). Although neither model reproduces actual observations with great precision, the inclusion of ship emissions does improve model performance—modestly, but with statistical significance.

Although including ship emissions slightly improves the overall performance of the model, results for specific regions vary. In particular, the model generally overpredicts SO_2 concentrations where DMS is the main source of sulphur. This can be seen in Fig. 3; we note the overpredictions in locations where the effect of ships is negligible and DMS emissions dominate (that is, data points for the two model runs are collocated). We believe that this may be related to the model's relatively high DMS to SO_2 yield. Several investigators^{11,16–18} have suggested that the DMS to SO_2 yield should be much lower than predicted by parametrizations developed to match observed marine atmospheric SO_2 concentrations^{6,7}. Their evidence is based on analysis of kinetic rates determined in the laboratory and observed dynamics of the DMS/ SO_2 system over the ocean. Capaldo and Pandis⁸ compared several of these DMS oxidation schemes, and determined that the low SO_2 yield (<0.1) of Hertel *et al.*¹⁷ produced significant underpredictions when compared to remote marine observations, and that the parametrizations of Pham *et al.*⁶ and Benkovitz *et al.*⁷ (SO_2 yields between 0.8 and 1.0) reproduced observations better (though even these high yields generated underpredictions of the SO_2 observations in the Northern Hemisphere oceans). Surhe *et al.*¹⁹ also found better agreement with observations near Tasmania using a higher DMS to SO_2 yield (0.8) compared to the lower value of De Bryun *et al.*¹¹ (0.3 to 0.5).

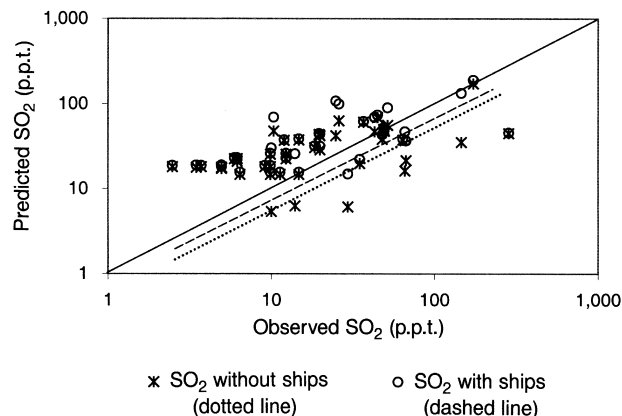


Figure 3 Comparison of observed and modelled SO_2 concentrations. Model predictions with ship emissions (circles) and without ship emissions (crosses), compared to the observed SO_2 concentrations for marine sites with observed SO_2 concentrations less than 300 p.p.t.v. The solid diagonal line represents the perfect predictive model while the additional lines are linear regressions (intercept forced through zero) for the two models: with ship emissions (dashed line; $r = 0.55$) and without ship emissions (dotted line; $r = 0.51$).

However, in a study of marine air off the coast of the state of Washington, Thornton *et al.*¹⁵ observed only a weak correlation between SO_2 and DMS concentrations¹⁵. These contradictions may be explained by the influence of ship emissions that, if unaccounted for, could bias the estimated DMS to SO_2 yield towards a higher value. Pham *et al.*⁶ and Benkovitz *et al.*⁷ used north Atlantic measurements to verify their yields, and Thornton *et al.*¹⁵ sampled air in the northeast Pacific. Figure 1 shows that these areas are influenced by ships. Although Surhe *et al.*¹⁹ and De Bryun *et al.*¹¹ used a location in the Southern Hemisphere considered to be more remote, the influence of shipping around Australia is still significant (Fig. 1).

In addition to these effects on ambient SO_2 and sulphate concentrations over the remote oceans, several continental areas also show a significant effect from ship emissions. This is expected, because nearly 70% of ocean-going ship emissions occur within 400 km of land³. As shown in Fig. 1, ship emissions often contribute more than 5% and as much as 30% to the modelled SO_2 concentrations in coastal regions. Ship emissions can also contribute 5% to 20% of the predicted NSS-sulphate concentrations over land (Fig. 2). Notable areas include parts of the west coasts of Canada and the United States, western Europe, Scandinavia, Japan, Indonesia, and the west coast of Africa. The effects of ship emissions on land are greater during the summer months, when land-based sulphur emissions in the Northern Hemisphere are lower, and may be twice as high as effects estimated by ship-emission studies limited to US port areas^{20–22}. These results for land regions demonstrate the importance of including ocean-going-ship emissions in regional air-quality studies.

Ship exhaust has been shown to increase local marine cloud albedo by adding to the available nuclei upon which cloud drops form^{23–25}. In addition, the large-scale effect of ship emissions of

Table 1 First-order sensitivity of the indirect forcing from ships

Global average indirect forcing	W m ⁻²
Base calculation	–0.11
Background CCN of 200 cm ⁻³ (from 100 cm ⁻³)	–0.06
CCN critical diameter $D_p^* = 0.08 \mu\text{m}$ (from $D_p^* = 0.10 \mu\text{m}$)	–0.21
4 day lifetime of PM, $V_{\text{dep}} = 0.3 \text{ cm s}^{-1}$ (from $V_{\text{dep}} = 0.4 \text{ cm s}^{-1}$)	–0.14
Continental influence cut-off of 0.5 p.p.b.v. NSS-sulphate (from 0.2 p.p.b.v.)	–0.16
CCN, cloud condensation nuclei; PM, particulate matter; V_{dep} , average CCN removal velocity; NSS, non-seasalt sulphate.	

particulates on the background concentration of CCN over the ocean could also have an effect on regional and global radiative budgets. To estimate this effect, we apply the fuel-based emission factors for particulate matter reported by Lloyds²⁶ to generate total annual particulate-matter emissions for ships (0.85 Tg yr⁻¹). This is then globally distributed according to the method of Corbett and Fischbeck^{2,3} to give us the emission rate of particulate matter ($E_{x,y}$) for any grid cell (x, y). We can then estimate the change in the CCN number, $\Delta\text{CCN}_{x,y}$, assuming steady state for the CCN concentration and using

$$\Delta\text{CCN}_{x,y} = \frac{E_{x,y}}{v_{\text{dep}}} f_m (D_p^*) N (D_p^*) \quad (1)$$

where v_{dep} is the average CCN removal velocity (wet and dry), f_m is the CCN mass fraction of the emitted particulate matter, N is the number density of CCN (number per mass larger than D_p^*) emitted by ships, and D_p^* is minimum diameter at which these particles activate to form cloud drops. Based on the particulate-matter size-distribution measurements of Lyyranen *et al.*²⁷, for large engines operating at full power, and assuming a D_p^* of 0.1 μm , we estimate f_m to be 0.6 (CCN per g particulate matter emitted) and N to be 10¹⁴ CCN per g of CCN. In addition, we assume a CCN lifetime of three days ($v_{\text{dep}} = 0.4 \text{ cm s}^{-1}$). Assuming as a first-order approximation that the change in cloud droplet number is equal to ΔCCN for the low CCN concentrations of the marine environment (100 cm⁻³), the potential effect of ships on annual average cloud albedo and radiative forcing can be estimated²⁸. To avoid problems with the nonlinear response of cloud droplet number to CCN concentration, and to keep our radiative forcing estimate conservative, we only consider ship particulate-matter emissions where continental influence is small. This is determined using our annual average model predictions for NSS-sulphate (with ships) and a threshold of 0.2 p.p.b.v. NSS-sulphate.

Using this method, we estimate the change in global radiative forcing due to cloud effects from ship particulate-matter emissions to be -0.11 W m^{-2} . This value is 14% of the IPCC estimate for 1990 global indirect forcing from all anthropogenic sulphate²⁹. The predicted average change in radiative forcing due to ships for the Northern Hemisphere is -0.16 W m^{-2} and for the Southern Hemisphere is -0.06 W m^{-2} . The direct effect of ship sulphur emissions is expected to be less dramatic²⁸. The sensitivity of our estimate to various uncertain input parameters is shown in Table 1. As can be seen, reasonable variation in individual input parameters can affect this first-order global estimate by at least a factor of 2.

Our results suggest that the emissions of sulphur and particulate matter from the international shipping industry need to be considered in the study of marine and coastal atmospheres. Because ship emissions, as a source of background sulphur, have been neglected in the past, many observational studies of the marine atmosphere need to be re-evaluated, particularly those in the remote oceans of the Northern Hemisphere. □

Received 19 February; accepted 5 July 1999.

1. Charlson, R. J., Lovelock, J. E., Andreae, M. O. & Warren, S. G. Oceanic phytoplankton, atmospheric sulphur, cloud albedo, and climate. *Nature* **336**, 655–661 (1987).
2. Corbett, J. J. & Fischbeck, P. S. Emissions from ships. *Science* **278**, 823–824 (1997).
3. Corbett, J. J., Fischbeck, P. S. & Pandis, S. N. Global nitrogen and sulfur emissions inventories for oceangoing ships. *J. Geophys. Res.* **104**, 3457–3470 (1999).
4. Benkovitz, C. M. *et al.* Global gridded inventories of anthropogenic emissions of sulfur and nitrogen. *J. Geophys. Res.* **101**, 29239–29253 (1996).
5. Kasibhatla, P., Chameides, W. L. & St John, J. A three-dimensional global model investigation of seasonal variations in the atmospheric burden of anthropogenic sulfate aerosols. *J. Geophys. Res.* **102**, 3737–3759 (1997).
6. Pham, M., Muller, J.-F., Brasseur, G. P., Granier, C. & Mégie, G. A three-dimensional study of the tropospheric sulfur cycle. *J. Geophys. Res.* **100**, 26061–26092 (1995).
7. Benkovitz, C. M. *et al.* Sulfate over the North Atlantic and adjacent continental regions: evaluation for October and November 1986 using a three-dimensional model driven by observation-derived meteorology. *J. Geophys. Res.* **99**, 20725–20756 (1994).
8. Capaldo, K. & Pandis, S. Dimethylsulfide chemistry in the remote marine atmosphere: Evaluation and sensitivity analysis of available mechanisms. *J. Geophys. Res.* **102**, 23251–23267 (1997).
9. Davison, B. & Hewitt, C. N. Elucidation of the troposphere reactions of biogenic sulfur species from a field measurement campaign in NW Scotland. *Chemosphere* **28**, 543–557 (1994).

10. Davison, B. *et al.* Dimethyl sulfide, methane sulfonic acid and physiochemical aerosol properties in Atlantic air from the United Kingdom in Halley Bay. *J. Geophys. Res.* **101**, 22855–22867 (1996).
11. De Bruyn, W. J., Bates, T. S., Caine, J. M. & Saltzman, E. S. Shipboard measurements of dimethylsulfide and SO₂ southwest of Tasmania during the first Aerosol Characterization Experiment (ACE 1). *J. Geophys. Res.* **103**, 16703–16711 (1998).
12. Ferek, R. J. *et al.* Dimethyl sulfide in the Arctic atmosphere. *J. Geophys. Res.* **100**, 26093–26104 (1995).
13. Pio, C. A., Cerqueira, M. A., Castro, L. M. & Salgueiro, M. L. Sulphur and nitrogen compounds in variable marine/continental air masses at the southwest European coast. *Atmos. Environ.* **30**, 3115–3127 (1996).
14. Talbot, R. W. *et al.* Chemical characteristics of continental outflow from Asia to the troposphere over the western Pacific Ocean during February–March 1994: Results from PEM-West B. *J. Geophys. Res.* **102**, 28255–28274 (1997).
15. Thornton, D. C. & Bandy, A. R. Sulfur dioxide and dimethyl sulfide in the central Pacific troposphere. *J. Atmos. Chem.* **17**, 1–13 (1993).
16. Bandy, A. R., Scott, D. L., Blomquist, B. W., Chen, S. M. & Thornton, D. C. Low yields of SO₂ from dimethyl sulfide oxidation in the marine boundary layer. *Geophys. Res. Lett.* **19**, 1125–1127 (1992).
17. Hertel, O., Christensen, J. & Hov, O. Modelling of the end products of the chemical decomposition of DMS in the marine boundary layer. *Atmos. Environ.* **28**, 2431–2449 (1994).
18. Saltelli, A. & Hjorth, J. Uncertainty and sensitivity analyses of OH-initiated dimethyl sulphide (DMS) oxidation kinetics. *J. Atmos. Chem.* **21**, 187–221 (1995).
19. Suhre, K. *et al.* Physico-chemical modeling of the first Aerosol Characterization Experiment (ACE 1) Lagrangian B1: A moving column approach. *J. Geophys. Res.* **103**, 16433–16455 (1998).
20. Carlton, J. S., Wright, A. A. & Coker, R. J. *Marine Exhaust Emissions—A Regional Survey of the English Channel* (Marine Management (Holdings) Ltd, London, 1994).
21. Port of Los Angeles, Port of Long Beach, Dames & Moore, Morrison and Foerster *Control of Ship Emission in the South Coast Air Basin: Assessment of the Proposed Federal Implementation Plan Ship Fee Emission Fee Program* (Port of Los Angeles, Los Angeles, California, 1994).
22. *Nonroad Engine and Vehicle Emission Study* (US Environmental Protection Agency, Washington DC, 1991).
23. Radke, L. F., Coakley, J. A. Jr & King, M. D. Direct and remote sensing observations of the effects of ships on clouds. *Science* **246**, 1146–1149 (1989).
24. King, M. D., Radke, L. F. & Hobbs, P. V. Optical properties of marine stratocumulus clouds modified by ships. *J. Geophys. Res.* **98**, 2729–2739 (1993).
25. Ferek, R. J., Hegg, D. A., Hobbs, P. V., Durkee, P. & Nielsen, K. Measurements of ship-induced tracks in clouds off the Washington coast. *J. Geophys. Res.* **103**, 23199–23206 (1998).
26. Carlton, J. S. *et al.* *Marine Exhaust Emissions Research Programme* (Lloyd's Register Engineering Services, London, 1995).
27. Lyyranen, J., Jokiniemi, J., Kauppinen, E., Joutsensaari, J. & Auvinen, A. Particle formation in medium speed diesel engines operating with heavy fuel oils. *J. Aerosol Sci.* **29**, S1003–S1004 (1998).
28. Seinfeld, J. H. & Pandis, S. N. *Atmospheric Chemistry and Physics: From Air Pollution to Climate Change* (Wiley & Sons, New York, 1998).
29. IPCC *Radiative Forcing of Climate Change. The 1994 Report of the Scientific Assessment Working Group of the Intergovernmental Panel on Climate Change (IPCC)* (Cambridge Univ. Press, 1995).

Acknowledgements. This work was supported by the US NSF and the NOAA Office of Global Programs.

Correspondence and requests for materials should be addressed to S.N.P. (e-mail: spyros@andrew.cmu.edu).

A pipiscid-like fossil from the Lower Cambrian of south China

D. Shu*, S. Conway Morris†, X.-L. Zhang*, L. Chen*, Y. Li* & J. Han*

* Department of Geology, Northwest University, Xi'an, 710069, People's Republic of China

† Department of Earth Sciences, University of Cambridge, Downing Street, Cambridge CB2 3EQ, UK

Exceptional fossil preservation is critical to our understanding of early metazoan evolution. A key source of information is the Burgess Shale-type faunas^{1–5}. Fossils from these deposits provide important insights into metazoan phylogeny, notably that of stem-group protostomes^{2,3,6}, and related topics such as trophic specialization⁷. Metazoan relationships are also being significantly reappraised in terms of molecular-based phylogenies^{8,9}, but integration of these data with palaeontological systematics is not straightforward^{10,11}. Moreover, molecular phylogenies are silent concerning the anatomies of stem-groups and the functional transitions that underpin the origin of different body plans^{2,6}. Some hitherto enigmatic fossils possess unique character-state combinations that, although they can be shoe-horned into extinct phyla¹², may be more profitably interpreted as defining major stem-groups^{2,3}. Here we describe a possible pipiscid, a metazoan previously known only from the Upper Carboniferous^{13,14}, from the Lower Cambrian of south China. Pipiscids

are currently interpreted as being agnathan chordates^{13–15}, but this discovery from the Chengjiang fossil-Lagerstätte indicates that the assignment of pipiscids to the Agnatha deserves to be reconsidered.

Phylum Uncertain

Xidazoon Shu, Conway Morris & Zhang gen. nov.

Xidazoon stephanus Shu, Conway Morris & Zhang sp. nov.

Etymology. Genus name an abbreviation of Chinese name for Northwest University at Xi'an. Species name *stephanos* (Greek) for crown.

Holotype. Early Life Institute, Northwest University, Xi'an. ELI-0000194.

Stratigraphy and locality. Qiongzhusi (Chiungchussu) Formation, Yu'an-shan member (*Eoredlichia* Zone); Lower Cambrian. Specimen collected from Haikou, Kunming, located about 50 km west of Chengjiang.

Diagnosis. Body with two-fold division, reminiscent of *Banffia* but anterior section more inflated and possessing prominent mouth circler. Anterior section with faint transverse divisions towards front, otherwise smooth. Mouth defined by circler of about 25 plates, divided into inner and outer regions, otherwise unarmed. Circler similar to plated mouth of *Pipiscius*, although in the latter taxon the plates are more cuticularized and inner circler folded into pharynx. Posterior section tapering towards front and back, segmented with cuticularized region of about six segments succeeded anteriorly by about three less well-defined segments. Posterior section similar to arthropodan metamerites, but lacking evidence of appendages. Cuticular segments also reminiscent of posterior region in *Yunnanozoon*, but in latter taxon segments are ventrally incomplete. Short terminal spines at posterior tip. Alimentary canal with terminal openings, anterior region possibly expanded and rectum with ?dilator muscles.

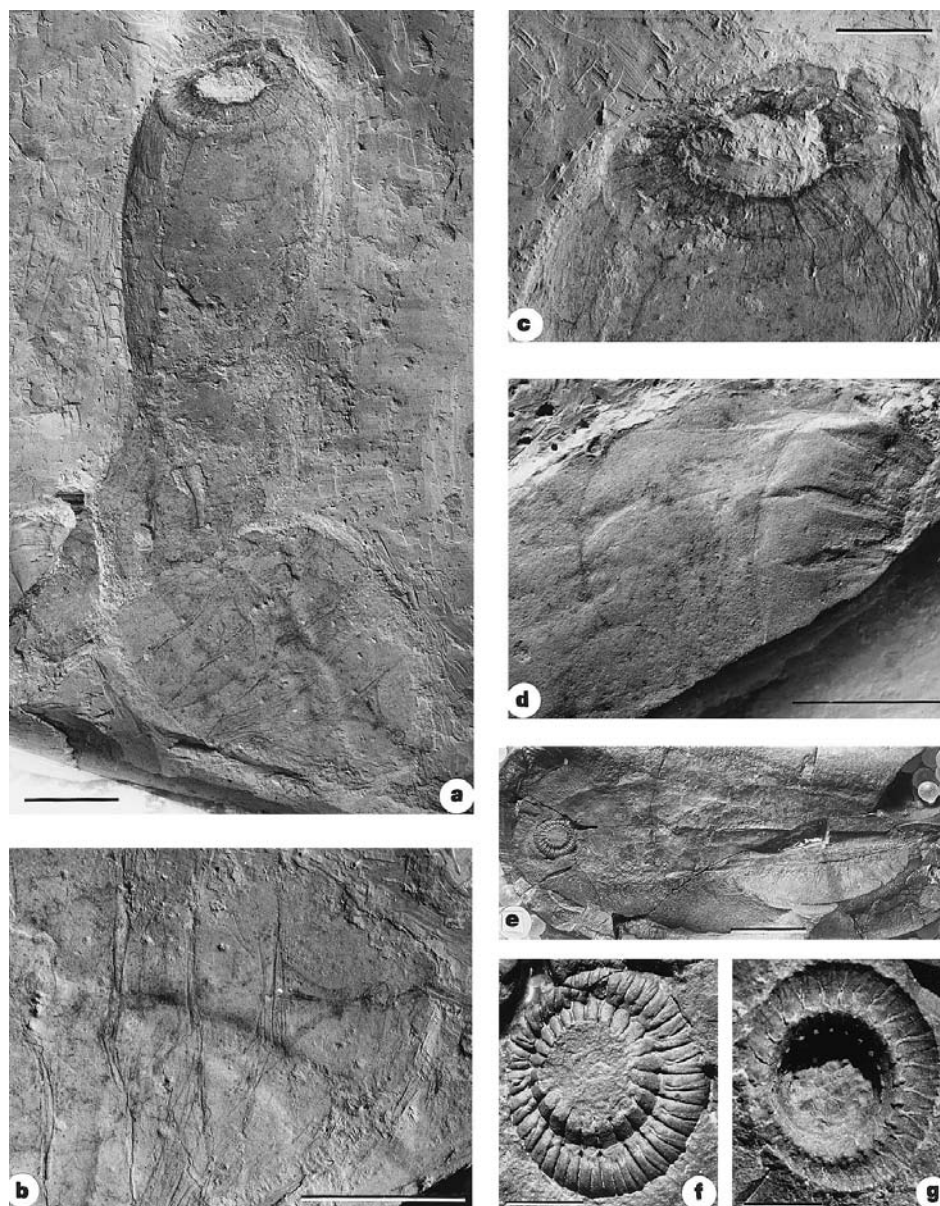


Figure 1 The Cambrian fossil *Xidazoon stephanus*, new species and Carboniferous ?agnathan *Pipiscius zangerli*. **a**, Entire specimen and (to lower left) incomplete individual of *Xidazoon* (compare to Fig. 2); **b**, detail of posterior section showing segmental divisions, gut trace, ?dilator muscles and posterior spines (right-hand side); **c**, detail of feeding apparatus of complete specimen; **d**, detail

of anterior and incomplete feeding apparatus of second specimen. **e**, Entire specimen of holotype of *Pipiscius*, part (PF 8345); **f**, detail of feeding apparatus of part; **g**, detail of feeding apparatus of counterpart. Scale bars: 10 mm (**a**, **b**, **e**), 5 mm (**c**, **d**) and 2 mm (**f**, **g**).

Description. *Xidazoon stephanus*, new genus and species, is known from two, or possibly three, specimens on a single slab (Figs 1a, 2). The most complete specimen is about 8.5 cm long and a second individual shows details of the anterior (Fig. 1d). The body comprises two main regions. The anterior section is moderately inflated, and the prominent circlet of the presumed anterior is interpreted as a feeding apparatus surrounding a voluminous mouth (Fig. 1c). The apparatus itself consists of plate-like structures, transversely folded to define inner and outer circlets. The edges of the inner circlet of plates are ridged (Fig. 1c), but they do not bear teeth or other extensions. In the second specimen (Fig. 1d) the plates appear to be separated anteriorly by narrower recessed areas. These may represent flexible inter-plate membranes. The anterior of the second specimen is incomplete, and that of the first is too crushed to give more than an estimate of the total number of plates. The better-preserved half-circumference displays about 13 plates, and an allocation of typical plate width around the circumference (~45 mm) gives a total of about 25. The mouth is gaping, but apart from the circlet of plates lacks evidence of jaws or other associated structures. Behind the feeding apparatus the anterior region bears faint, widely separated transverse divisions that may be segmental.

The posterior section tapers in either direction from an expanded central zone. It consists of about six well-defined segments (Figs 1b, 2), and in the anterior direction there is a series of more faint transverse annulations. The surface appears to have been lightly cuticularized. The segment boundaries vary from tightly adpressed to separated, indicating originally relatively wide and flexible inter-segmental membranes. At the posterior tip there are two or three spinose projections (Figs 1b, 2).

Little is known of the internal anatomy. A gut trace is present in the mid and posterior sections, and near the terminal anus diverging strands may represent dilator muscles (Figs 1a, b, 2). Towards the anterior of the visible gut trace it appears to expand, and in the anterior section it may have been voluminous.

Preservation. The style and quality of preservation is similar to other Chengjiang taxa, such as *Yunnanozoon*^{16–19}. Thus, the extent of decay appears to be limited. Features, notably the circlet of plates and the posterior segmentation, seem to be original rather than post-mortem artefacts.

Ecology. The ecology of *Xidazoon* is problematic. It was presumably benthic, with the anterior circlet periodically contracting to ingest detritus. The inflated nature of the anterior section in the most complete specimen might be because of sediment ingestion. An alternative possibility is that the anterior organ acted as a sucker for lodgement on prey or hard substrates.

Discussion. Comparisons between *Xidazoon* and various extant metazoan groups, such as the sipunculans and the much smaller cyclophorans, are not convincing. Similarly, among the diverse Burgess Shale-type assemblages, no exact counterpart to *Xidazoon* has been recognized. There are some similarities to the otherwise enigmatic *Banffia confusa*⁵, which consists of a segmented unit, apparently posterior to an elongate and smooth section, but this taxon lacks evidence for the prominent feeding apparatus of *Xidazoon*. The better-known anomalocaridids²⁰ have a prominent circular feeding apparatus and a bipartite body with segmented posterior section. There are, however, many differences. The feeding apparatus occurs in a variety of forms^{3,5,20}, but none is particularly similar to *Xidazoon*. Other characteristic features of the anomalocaridids, notably the anterior giant appendages and lateral lobes, have no parallel in this new fossil.

The anterior circlet of *Xidazoon* is, however, similar to the otherwise unique feeding apparatus of the putative agnathan *Pipiscius zangerli* (Fig. 1e), a rare species from the 300-Myr-old Mazon Creek fossil-Lagerstätte (Upper Carboniferous) of Illinois. The original description¹³ is convoluted, but in essence the feeding apparatus is composed of two circles of sclerotized plates. The inner series ('collar lamellae' of ref. 13) total 23, a number with no apparent parallel in other metazoan organ systems. The outer circlet is also cited¹³ as consisting of 23 plates. There is, however, a hitherto

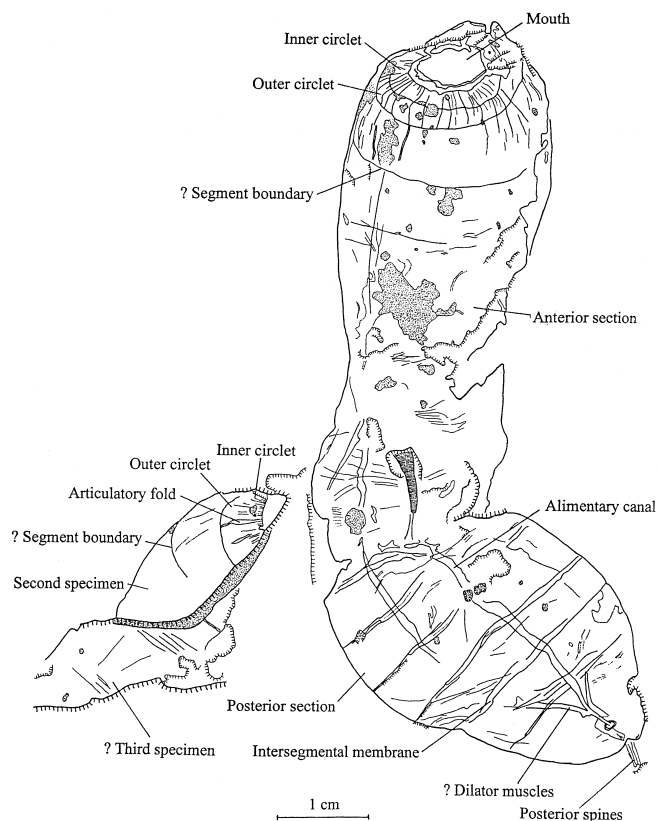


Figure 2 Camera-lucida drawing of slab containing the two (or possibly three) specimens of the Cambrian fossil *Xidazoon stephanus*, new species.

unrecognized duplication on the leading anterior plate, so that the total number of plates is effectively 24. This duplication defines a line of bilateral symmetry in the apparatus. The plates are separated by narrow clefts ('vanes' of ref. 13) that presumably accommodated shape changes associated with feeding. The principal similarities between the anterior apparatus of *Pipiscius* and *Xidazoon* are the double nature of the circlet with direct continuity between the inner and outer plates, the similar number of plates and evidence for articulatory zones (Fig. 1d) that seem to be comparable to the 'vanes' (Fig. 1f). The apparatus, however, are not identical. In *Pipiscius* the outer plates have a more complex structure, housing triangular insets. These latter units may have accommodated movement of the apparatus, possibly necessitated by a more pronounced sclerotization. Deep pits associated with the 'vanes', and possibly employed for muscle insertions¹³, are not evident in *Xidazoon*. Finally, the inner circle ('collar') of *Pipiscius* is directed inwards, whereas in *Xidazoon* it appears to be more rim-like.

Concerning the possible connection between *Xidazoon* and *Pipiscius*, there seems to be three alternative evolutionary scenarios. First, the annular feeding apparatus is simply an example of convergence. Among the many suctorial and other biological attachment structures similarities can be shown, for example, with the attachment organ of the ectoparasitic ciliate *Trichodina pediculus*²¹ and the arm suckers of the octopus²², although no phylogenetic connection with *Xidazoon* can be seriously entertained. Notwithstanding the bi-annular arrangement of about 25 plates, the few similarities that otherwise exist between *Xidazoon* and *Pipiscius* make convergence a reasonable option. Second, *Xidazoon* and *Pipiscius* are related, but the assignment of the latter taxon to the agnathans^{13,14} is erroneous: together they would represent a new major Palaeozoic clade of as yet unknown affinities. In this sense it would be comparable to such enigmatic groups as the typhloesids²³ and tullimonstrids²⁴.

The third proposal is that *Xidazoon* is a precursor to the agnathans, including *Pipiscius*. This presupposes the homology of the circular feeding apparatus in the two taxa, and that certain features (such as fin-rays and possible myotomes) of *Pipiscius* are indicative of a chordate relationship. In this scenario *Xidazoon* would potentially provide new insights into the organization of stem-group deuterostomes. A link may also exist with the coeval *Yunnanozoon*^{16–19}. This Chengjiang taxon displays putative gill slits, and the cuticular segmentation has some similarities with *Xidazoon*. One reconstruction¹⁸ of *Yunnanozoon* also depicts a circum-oral set of plates. The bipartite nature of *Xidazoon* is more strongly developed than in *Yunnanozoon*, but the almost arthropod-like segmented posterior section could provide an intriguing phylogenetic link with the protostomes²⁵. Continuing investigations of Lower Cambrian fossil-Lagerstätten may yield relatives of *Xidazoon* that will help to resolve the controversial status of these fossils in the context of metazoan phylogeny. □

Received 19 November 1998; accepted 2 July 1999

1. Conway Morris, S. *The Crucible of Creation: The Burgess Shale and the Rise of Animals* (Oxford Univ. Press, Oxford, 1998).
2. Conway Morris, S. & Peel, J. S. Articulated halkieriids from the Lower Cambrian of North Greenland and their role in early protostome evolution. *Phil. Trans. R. Soc. Lond. B* **347**, 305–358 (1995).
3. Budd, G. E. in *Arthropod Relationships* (eds Fortey, R. A. & Thomas, R. H.) *Syst. Ass. Spec. Vol.* **55**, 125–138 (1997).
4. Chen, J.-Y. et al. *The Chengjiang Biota* (National Museum of Natural Science, Taiwan, c. 1996).
5. Chen, J.-Y. & Zhou, G.-Q. Biology of the Chengjiang fauna. *Bull. Natl Mus. Nat. Sci. Taiwan* **10**, 11–105 (1997).
6. Budd, G. E. The morphology of *Opabinia regalis* and the reconstruction of the arthropod stem-group. *Lethaia* **29**, 1–14 (1996).
7. Butterfield, N. J. Burgess Shale-type fossils from a Lower Cambrian shallow-shelf sequence in northwestern Canada. *Nature* **369**, 477–479 (1994).
8. de Rosa, R. et al. Hox genes in brachiopods and priapulids and protostome evolution. *Nature* **399**, 772–776 (1999).
9. Ruiz-Trillo, I. et al. Acoel flatworms: Earliest extant bilaterian metazoans, not members of platyhelminthes. *Science* **283**, 1919–1923 (1999).
10. Conway Morris, S. Why molecular biology needs palaeontology. *Development* (Suppl.) **1994**, 1–13 (1994).
11. Conway Morris, S. Metazoan phylogenies: falling into place or falling to pieces? A palaeontological perspective. *Curr. Op. Genet. Dev.* **8**, 662–667 (1998).

12. Gould, S. J. *Wonderful Life: The Burgess Shale and the Nature of History* (Norton, New York, 1989).
13. Bardack, D. & Richardson, E. S. New agnathous fishes from the Pennsylvanian of Illinois. *Fieldiana Geol.* **33**, 489–510 (1977).
14. Bardack, D. in *Richardson's Guide to the Fossil Fauna of Mazon Creek* (eds Shabica, C. W. & Hay, A. A.) 226–243 (Northeastern Illinois Univ. Press, Chicago, 1997).
15. Janvier, P. *Early Vertebrates* (Clarendon, Oxford, 1996).
16. Chen, J.-Y. et al. A possible early Cambrian chordate. *Nature* **377**, 720–722 (1995).
17. Chen, J.-Y. & Li, C.-W. Early Cambrian chordate from Chengjiang, China. *Bull. Natl Mus. Nat. Sci. Taiwan* **10**, 257–273 (1997).
18. Dzik, J. *Yunnanozoon* and the ancestry of chordates. *Acta Palaeont. Pol.* **40**, 341–360 (1995).
19. Shu, D., Zhang, X.-L. & Chen, L. Reinterpretation of *Yunnanozoon* as the earliest known hemichordate. *Nature* **380**, 428–430 (1996).
20. Collins, D. The "evolution" of *Anomalocaris* and its classification in the arthropod class Dinocarida (nov.) and order Radiodonta (nov.). *J. Paleont.* **70**, 280–293 (1996).
21. Nachtigall, W. *Biological Mechanisms of Attachment* (Springer, Berlin, 1974).
22. Packard, A. in *The Mollusca, Form and Function Vol. 11* (eds Trueman, E. R. & Clarke, M. R.) 37–67 (Academic, San Diego, 1988).
23. Conway Morris, S. *Typhloesus wellsi* (Melton and Scott, 1973), a bizarre metazoan from the Carboniferous of Montana, USA. *Phil. Trans. R. Soc. Lond. B* **327**, 595–624 (1990).
24. Johnson, R. G. & Richardson, E. S. Pennsylvanian invertebrates of the Mazon Creek area, Illinois: The morphology and affinities of *Tullimonstrum*. *Fieldiana Geol.* **12**, 119–149 (1969).
25. Holland, L. Z. & Holland, N. D. Developmental gene expression in *Amphioxus*: New insights into the evolutionary origin of vertebrate brain regions, neural crest, and rostral/caudal segmentation. *Am. Zool.* **38**, 647–658 (1998).

Acknowledgements. We thank the National Foundation of Natural Sciences of China, Minister of Science and Technology of China, Royal Society, National Geographic Society, and St John's College, Cambridge for support. Access to *Pipiscius* was facilitated by P. Crane (Field Museum, Chicago), and M. P. Smith and P. Donoghue (University of Birmingham). Technical assistance by S. J. Last and D. R. Simons is acknowledged, as are critical comments by P. Janvier, D. B. Norman and S. Jensen.

Correspondence and requests for materials should be addressed to D.S. (e-mail: dgshu@sein.sxgb.com.cn).

Environmental controls on the geographic distribution of zooplankton diversity

Scott Rutherford*, Steven D'Hondt* & Warren Prell†

* Graduate School of Oceanography, University of Rhode Island, Narragansett, Rhode Island 02882, USA

† Geological Sciences, Brown University, Providence, Rhode Island 02912, USA

Proposed explanations for the geographic distribution of zooplankton diversity include control of diversity by geographic variation in: physical and chemical properties of the near-surface ocean^{1–3}; the surface area of biotic provinces⁴; energy availability⁵; rates of evolution and extinction⁶; and primary productivity⁷. None of these explanations has been quantitatively tested on a basin-wide scale. Here we used assemblages of planktic foraminifera from surface sediments to test these hypotheses. Our analysis shows that sea-surface temperature measured by satellite⁸ explains nearly 90% of the geographic variation in planktic foraminiferal diversity throughout the Atlantic Ocean. Temperatures at depths of 50, 100 and 150 m (ref. 9) are highly correlated to sea-surface temperature and explain the diversity pattern nearly as well. These findings indicate that geographic variation in zooplankton diversity may be directly controlled by the physical structure of the near-surface ocean. Furthermore, our results show that planktic foraminiferal diversity does not strictly adhere to the model of continually decreasing diversity from equator to pole. Instead, planktic foraminiferal diversity peaks in the middle latitudes in all oceans.

We used the Brown University Foraminiferal Data Base (BFD, 33 species and 6 subspecies >150 µm; 1,252 samples) to document the global pattern of planktic foraminiferal diversity and to evaluate the long-standing hypothesis of a latitudinal diversity gradient. We used planktic foraminifera for this study because all of the living species are known and are included in the BFD, unlike other plankton groups (such as radiolaria or copepods), which have hundreds of species and greater taxonomic uncertainty. Our results clearly show that, throughout the world ocean, planktic foraminiferal diversity peaks in middle latitudes, is lowest at high latitudes and is inter-

Table 1 Summary of regression results for the North and South Atlantic

r^2	RMSE	No. of variables	Variables in model		Type of fit (var1/var2)
0.91	2.50	2	Mean annual SST ⁸	Temperature at 150 m	3rd order polynomial/3rd order polynomial
0.90	2.68	2	Mean annual SST ⁸	Mean annual chlorophyll-a ²⁴	3rd order polynomial/exponential
0.89	2.71	2	Mean annual SST ⁸	Salinity ²⁵	3rd order polynomial/linear
0.89	2.72	2	Mean annual SST ⁸	Dissolved nitrate ²⁵	3rd order polynomial/linear
0.89	2.74	1	Mean annual SST ⁸		3rd order polynomial
0.89	2.77	2	Annual solar irradiance ²⁴	Temperature at 150 m ⁹	Linear/3rd order polynomial
0.88	2.95	1	Temperature at 150 m ⁹		3rd order polynomial
0.81	3.64	1	Mean annual SST ⁸		Linear
0.81	3.70	1	Temperature at 150 m ⁹		Linear
0.77	4.04	1	Annual solar irradiance ²⁴		Linear
0.60	5.23	1	Mean annual chlorophyll-a ²⁴		Exponential
0.53	5.78	1	Salinity ²⁵		Linear
0.43	6.32	1	Mean annual topography ²⁶		Linear
0.10	7.97	1	Dissolved nitrate ²⁵		Linear

Third-order polynomial regressions of SST or temperature at 150 m depth provide the most parsimonious fit. Other variables investigated but not shown owing to poor correlations are mixed-layer depth²⁷, annual mixed-layer depth variance, annual chlorophyll variance, annual SST variance, seasonal SST difference, peak spring primary productivity²⁸, apparent oxygen utilization²⁹, the one percent light depth calculated from^{24,29} and mean geostrophic current velocity²⁸. $n = 351$.

mediate in the topics (Fig. 1). This result strongly contrasts with the common perception that planktic foraminiferal diversity decreases continually from equator to pole^{4,6}. This pattern of peak diversity in middle latitudes may not be limited to planktic foraminifera; euphausiids, pteropods and chaetognaths caught in North Pacific plankton tows show a similar pattern^{10,11}.

Although our results are presented as the number of species present in a sample (simple species richness), we have taken steps to ensure that the results are not affected by sample-size variation. The pattern of diversity does not change if we use the Margalef index¹² (which normalizes the number of species in a sample to the sample size), remove all samples with less than 250 individuals (19 samples), or remove rare species (those comprising less than 1% of a sample) from each sample. In addition, our bootstrap simulations using a highly dominated assemblage of 30 species show that species numbers in core-top foraminiferal assemblages are sufficiently low and evenly distributed that counts of a few hundred individuals are enough to identify nearly all of the species present in the assemblage. These simulations were based on statistical 'sub-sampling' of actual foraminiferal census data. The simulations show that a sample size of 450 individuals has a 99% probability of encountering a species that comprises 1% of the assemblage. The average number of individuals in our foraminiferal samples is 529.

We correlated planktic foraminiferal diversity to remotely-sensed and directly-measured environmental variables to investigate quantitatively how effectively these variables can predict the geographic

distribution of diversity. Because calcium carbonate dissolution is ubiquitous in deep-sea sediments of the Pacific and Indian Oceans and may decrease the species richness of core-top assemblages, we confined our statistical analysis to the Atlantic Ocean. To minimize the effect of dissolution on our Atlantic results, we deleted all samples taken from water depths greater than 4,500 m (the approximate depth of significant calcite dissolution in the North Atlantic¹³).

The best predictors of foraminiferal diversity are mean annual sea-surface temperature (SST)⁸ and mean annual temperature at depth (50, 100 or 150 m)⁹. A third-order polynomial equation including any one of these variables explains nearly 90% of the variability in species richness (or in the Margalef index of diversity) and the addition of other variables to the model does not significantly improve the fit (Table 1).

Maps of actual diversity and diversity calculated using the polynomial SST equation illustrate how well the geographic distribution of diversity is reproduced by the water temperature regressions (Fig. 2). This result indicates that satellite measurements of the skin temperature of the ocean (the upper few millimetres) can be used to predict zooplankton diversity measured by counts of skeletons that have accumulated in the sediments over decades to centuries. Furthermore, this relationship applies to the entire Atlantic Ocean. The fact that euphausiids, pteropods and chaetognaths from North Pacific plankton tows^{10,11} exhibit a geographic pattern of diversity similar to that of planktic foraminifera indicates that these correlations may apply to many plankton groups.

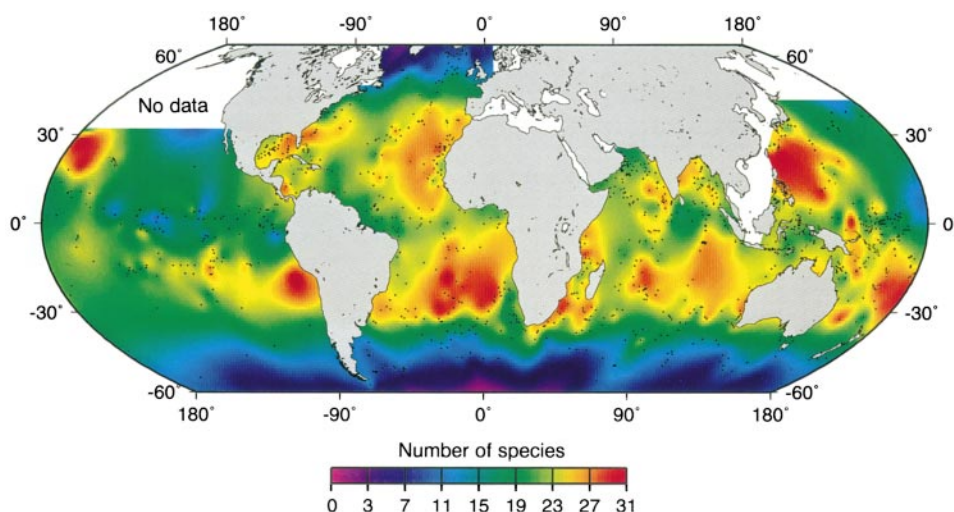


Figure 1 Global pattern of planktic foraminiferal diversity from 1,068 surface-sediment samples. All core-top samples from water depths less than 4,500 m were included. Note that peak diversity generally occurs in middle latitudes in

both the northern and southern hemispheres in all oceans. All colour contour plots were created using The Generic Mapping Tools utilities *blockmean* ($2^\circ \times 2^\circ$ grid), *surface* (tension, 0.25; convergence, 0.1) and *grdimage* ($1^\circ \times 1^\circ$ grid).

Table 2 Comparison of species diversity with area of biotic province

Area $\times 10^7$ (km ²)	Areally weighted diversity	Biotic province ¹⁵
0.697	23.8	temperate
0.862	13.7	subpolar
1.106	24.8	subtropical
1.258	20.8	tropical

Area of biotic provinces defined by Bogdanov¹⁵ compared with area-weighted diversity in each province. Area-weighted diversity is calculated by gridding the 351 diversity data points to a $1^\circ \times 1^\circ$ grid. Each grid box contains an average diversity that is multiplied by the area (km²) of the box and summed over the entire province. The total is then divided by the total area of the province.

The residuals from our polynomial regressions of water temperature to foraminiferal diversity are not completely random (Fig. 3). Positive residuals occur in the eastern South Atlantic (where temperature underpredicts diversity) and negative residuals are located in the western tropical Atlantic (where temperature overpredicts diversity). The region of overprediction in the western tropical Atlantic is centred on the Amazon and Orinoco outflows, indicating that salinity or turbidity may influence foraminiferal diversity in this area. However, these negative residuals are also consistent with enhanced dissolution at water depths shallower than 4,500 m.

The positive linear correlation between SST and temperature at 50, 100 or 150 m water depth for SSTs between -2 and 27°C ($r^2 > 0.94$ for each depth) implies that, in much of the world ocean SST and thermocline structure covary. This explains why residuals from the SST and temperature-at-depth polynomial regressions are so similar (Fig. 3). In the tropics, where SST is greater than 27°C , SST is not related to temperature at depth ($r^2 < 0.1$ for each depth), indicating that different factors, such as equatorial currents, control upper ocean structure in this region.

Although the various plankton diversity models introduced above^{1–7} have persisted in the literature for many years (see refs 4, 21 for reviews), none has been quantitatively tested over an entire ocean basin. We now explicitly test these models using our foraminiferal data and correlations to surface water properties. The species–area hypothesis predicts that larger biotic provinces will contain more species⁴. This model implicitly accepts that broad open-ocean biotic provinces are bounded by oceanographic fronts.

We used Bogdanov's¹⁵ definitions of oceanic provinces to test the species–area hypothesis. The geographic area encompassed by the biotic provinces and the area-weighted diversity of each region do not show the relationship predicted by the hypothesis (Table 2). The tropical province covers the largest area but has an intermediate diversity, and the temperate province, with one of the highest diversities, covers the smallest area. Thus we conclude that the area of a biotic province does not exert primary control on planktic foraminiferal diversity.

Solar energy has been widely cited as controlling diversity^{6,7}. On the basis of our regressions, incident solar radiation explains a smaller percentage of the diversity variation ($r^2 = 0.77$ for a linear regression and 0.82 for a third-order polynomial) than does SST (Table 1). At first glance, our correlations between SST and diversity might indicate that energy entering the ocean as sensible heat controls the geographic distribution of planktic foraminiferal diversity. However, the best correlation is not linear. For example, the polynomial SST relationship predicts that 23 species will inhabit water at temperatures of 20.6°C and 29.7°C , whereas 26 species will be present in water at 25°C . Thus, increasing energy does not continually increase the number of species inhabiting a region. The decreasing diversity at temperatures $>27^\circ\text{C}$ might be explained by the inability of planktic foraminifera to maintain physiological processes at these high temperatures. However, culturing experiments show that many species of planktic foraminifera can feed, grow and reproduce at temperatures over 30°C ¹⁶. Furthermore, many planktic foraminifera inhabit the base of the mixed layer and the thermocline, where temperatures are a few degrees cooler than at the surface.

Solar energy also enters ecosystems through primary production. Our results show that, relative to temperature, both chlorophyll concentration (mean annual, seasonal and peak) from the Coastal Zone Color Scanner (CZCS) and recent estimates of peak seasonal primary productivity are poor predictors of zooplankton diversity (Table 1).

It has been suggested that energy influences diversity by driving faster evolutionary rates in the tropics^{6,14}. However, rates of speciation, extinction and turnover in Neogene planktic foraminifera¹⁷ show no significant differences between tropical, transitional and

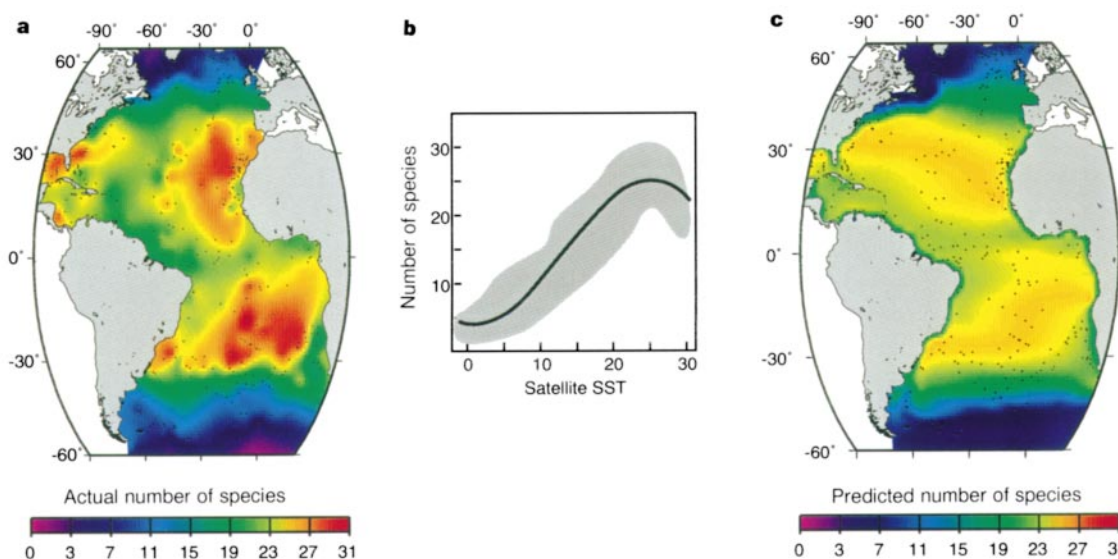


Figure 2 Actual and predicted planktic foraminiferal diversity in the Atlantic. **a**, Actual number of species based on 351 core-top samples. **b**, The nonlinear relationship between satellite SST and diversity (solid line) and the distribution of the data (shaded area). Note that there is a positive relationship between SST and diversity from -2°C to 27°C and a negative relationship above 27°C . **c**, Planktic

foraminiferal diversity predicted from SST using the third-order polynomial equation shown in **b**. The geographic pattern of diversity is reproduced very well except in a portion of the North Atlantic where there are no data points and in the western tropical Atlantic near the Amazon and Orinoco outflows.

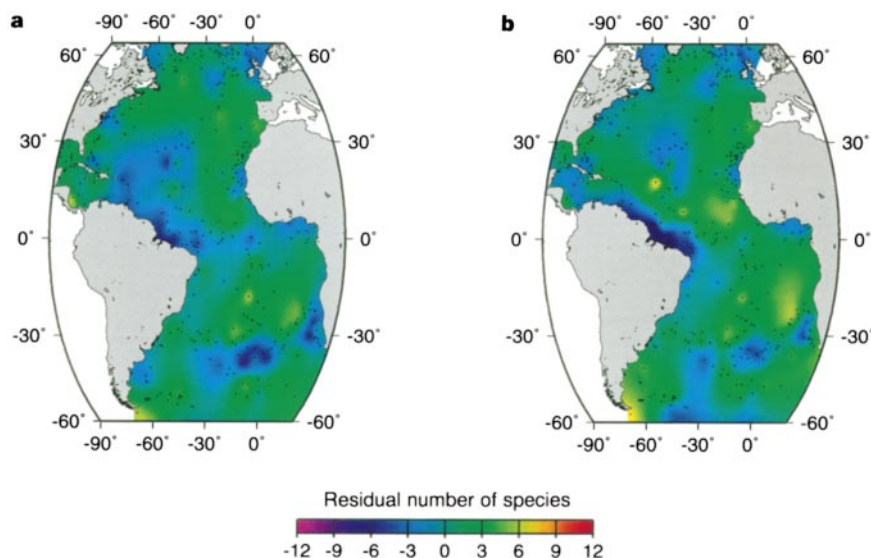


Figure 3 Residual diversity (actual diversity minus predicted diversity) from the SST polynomial regression (**a**) and the temperature at 150m water depth polynomial regression (**b**). In both cases negative residuals (where the predicted

diversity exceeds the actual diversity) are concentrated near the Amazon and Orinoco outflow. The similarity of the residuals indicates that the thermal structure of the water column may control species diversity.

temperate assemblages. Furthermore, peak planktic foraminiferal diversity (Fig. 1) and peak diversity of euphausiids, pteropods and chaetognaths occur in middle latitudes^{10,11}, not in the tropics where evolutionary rates are supposedly faster.

Our correlations of diversity to near-surface water temperature lead us to suggest that geographic variation in planktic foraminiferal diversity is primarily controlled by vertical niche partitioning dictated by the thermal structure of the near-surface ocean. A thermocline with a gradual temperature change and a deep base may provide more niches per unit surface area than a sharp thermocline with a shallow base. For example, some species (*Globigerinoides ruber* and *Globigerinoides sacculifer*) live in the mixed layer, whereas others live in the upper thermocline (*Neogloboquadrina dutertrei* and *Globorotalia menardii*), mid-thermocline (*Globorotalia scitula*) or deepest thermocline (*Globorotalia truncatulinoides* and *Globorotalia hirsuta*)¹⁸. Where a gradual thermocline is present, these species can inhabit the same region at different water depths.

Comparing the Atlantic planktic foraminiferal diversity pattern to near-surface temperature measurements and meridional cross-sections of potential density¹⁹ leads us to the following model of geographic variation in zooplankton diversity. In the high latitudes, where diversity is lowest, the thermocline is nearly absent and there is little vertical partitioning of niches. The middle latitudes are characterized by higher SST, a thick permanent thermocline and greatest diversity. In the tropics, diversity decreases where SST is greater than 27 °C because this region has a very sharp thermocline with a shallow base, and thus decreased vertical niche separation.

Our hypothesis that vertical niche separation maintains geographic variability in zooplankton diversity is consistent with studies conducted in the North Pacific. Plankton-tow studies of copepods²⁰ and phytoplankton²¹ indicated that groups of species occur as vertically segregated communities. A Doppler sonar study of organisms less than 2.0-cm long showed discrete scattering layers at 300, 560 and 1,000 m water depth²². Although those studies showed that vertical segregation of communities occurs locally, our results indicate that it may be a global phenomenon. Identifying the processes that might allow similar species to inhabit the same vertically segregated community remains a difficult problem²⁰, however, and our data do not allow us to assess these processes directly.

In closing, our correlations of planktic foraminiferal diversity to upper-ocean thermal structure provide direct evidence that the geographic pattern of planktic foraminiferal diversity results from geographic differences in the physical structure of the near-surface ocean. They indicate that first-order differences in the diversity of different biotic provinces may not require explanations that vary from one province to another (for example, mixing of faunas in one region, increased productivity in another). They indicate that the difference in diversity between the mid-latitude gyres and the equatorial ocean is not a simple function of surface area. The close relationship between water-column structure and ecosystem structure explains the success of foraminifer-based palaeo-SST estimation schemes. Just as bird diversity in a forest can be related to the number of vegetation levels²³, we suggest that zooplankton diversity is primarily controlled by vertical niche availability dictated by the thermal structure of the near-surface ocean. □

Received 8 March; accepted 18 June 1999.

- Ruddiman, W. F. Recent planktonic Foraminifera: Dominance and diversity in North Atlantic surface sediments. *Science* **164**, 1164–1167 (1969).
- Williams, D. F. & Johnson, W. C. Diversity of recent planktonic foraminifera in the Southern Indian Ocean and Late Pleistocene paleotemperatures. *Quat. Res.* **5**, 237–250 (1975).
- Balsam, W. L. & Flessa, K. W. Patterns of planktonic foraminiferal abundance and diversity in surface sediments of the western North Atlantic. *Mar. Micropaleontol.* **3**, 279–294 (1978).
- Rosenzweig, M. L. *Species Diversity in Space and Time* (Cambridge Univ. Press, Cambridge, 1995).
- Connell, J. H. & Orias, E. The ecological regulation of species diversity. *Am. Nat.* **98**, 399–414 (1964).
- Stehli, F. G., Douglas, R. G. & Newell, N. D. Generation and maintenance of gradients in taxonomic diversity. *Science* **164**, 947–949 (1969).
- Currie, D. J. Energy and large-scale patterns of animal and plant species richness. *Am. Nat.* **137**, 27–49 (1991).
- Casey, K. S. & Cornillon, P. A comparison of satellite and *in situ* based sea surface temperature climatologies. *J. Clim.* **12**, 1848–1863 (1999).
- Osborne, J. & Flinchem, E. P. *Ocean Atlas* (NOAA/PMEL, Seattle, 1994).
- Reid, J. L., Brinton, E., Fleminger, A., Venrick, E. L. & McGowan, J. A. in *Advances in Oceanography* (eds Charnock, H. & Deacon, G.) 65–130 (Plenum, New York, 1978).
- McGowan, J. A. & Walker, P. W. Dominance and diversity maintenance in an oceanic ecosystem. *Ecol. Monogr.* **55**, 103–118 (1985).
- Margalef, R. Information theory in ecology. *Gen. Syst.* **3**, 36–71 (1958).
- Pilson, M. E. Q. *An Introduction to the Chemistry of the Sea* (Prentice-Hall, Upper Saddle River, New Jersey, 1998).
- Rohde, K. Latitudinal gradients in species diversity: The search for the primary cause. *Oikos* **65**, 514–527 (1992).
- Bogdanov, D. V. Map of the natural zones of the ocean. *Okeanologiya* **1**, 941–943 (1961).
- Bijma, J., Faber, W. F. & Hemleben, C. Temperature and salinity limits for growth and survival of some planktonic foraminifera in laboratory culture. *J. Foramin. Res.* **20**, 95–116 (1990).
- Wei, K.-Y. & Kennett, J. P. Taxonomic evolution of Neogene planktonic foraminifera and paleoceanographic relations. *Paleoceanography* **1**, 67–84 (1986).
- Hemleben, C., Spindler, M. & Anderson, O. R. *Modern Planktonic Foraminifera* (Springer, New York, 1989).
- Bainbridge, A. E. *GEOSCS Atlantic Expedition 2* (National Science Foundation, Washington, 1982).

20. McGowan, J. A. & Walker, P. W. Structure in the copepod community of the North Pacific central gyre. *Ecol. Monogr.* **49**, 195–226 (1979).
21. Venrick, E. L. Phytoplankton in an oligotrophic ocean: Observations and questions. *Ecol. Monogr.* **52**, 129–154 (1982).
22. Plueddemann, A. J. & Pinkel, R. Characterization of the patterns of diel migration using a Doppler sonar. *Deep-Sea Res.* **36**, 509–530 (1989).
23. MacArthur, R. H. & MacArthur, J. On bird species diversity. *Ecology* **42**, 594–598 (1961).
24. SeaWiFS Project, G.S.F.C., Distributed Active Archive Center. <http://seawifs.gsfc.nasa.gov/SEAWIFS>; <http://www.giss.nasa.gov>.
25. Levitus, S. & Boyer, T. P. *World Ocean Atlas* (National Oceanic and Atmospheric Administration, Washington, 1994).
26. Halpern, D., Zlotnicki, V., Brown, O., Freilich, M. & Wentz, F. *An Atlas of Monthly Mean Distributions of SSM/I Surface Wind Speed, AVHRR/2 Sea Surface Temperature, AMI Surface Wind Velocity and TOPEX/POSEIDON Sea Surface Height During 1994* (Jet Propulsion Lab., Pasadena, California, 1995).
27. Levitus, S. *Climatological Atlas of the World Ocean* (National Oceanic and Atmospheric Administration, Rockville, MD, 1982).
28. Howard, K. L. *Estimating Global Ocean Primary Productivity Using Satellite-Derived Data* (Univ. Rhode Island, 1995).
29. Nelson, D. M. & Smith, W. O. Sverdrup revisited: critical depths, maximum chlorophyll levels, and the control of Southern Ocean productivity by the irradiance-mixing regime. *Limnol. Oceanogr.* **36**, 1650–1661 (1991).
30. Stammer, D. & Wunsch, C. Preliminary assessment of the accuracy and precision of TOPEX/POSEIDON altimeter data with respect to the large-scale ocean circulation. *J. Geophys. Res.* **99**, 24584–24604 (1994).

Acknowledgements. We thank J. Yoder and J. Ryan for providing the primary productivity and mixed layer data; P. Cornillon and K. Casey for providing the satellite SST data; D. Stammer for providing the geostrophic current velocity data; J. A. McGowan for thoughtful comments; and the U.R.I. Graduate School of Oceanography for publication support.

Correspondence and requests for materials should be addressed to S.R. (e-mail: rutherford@deschutes.gso.uri.edu).

Context generalization in *Drosophila* visual learning requires the mushroom bodies

Li Liu, Reinhard Wolf, Roman Ernst & Martin Heisenberg

Lehrstuhl für Genetik, Biozentrum, D-97074 Würzburg, Germany

The world is permanently changing. Laboratory experiments on learning and memory normally minimize this feature of reality, keeping all conditions except the conditioned and unconditioned stimuli as constant as possible¹. In the real world, however, animals need to extract from the universe of sensory signals the actual predictors of salient events by separating them from non-predictive stimuli (context²). In principle, this can be achieved if only those sensory inputs that resemble the reinforcer in

their temporal structure are taken as predictors. Here we study visual learning in the fly *Drosophila melanogaster*, using a flight simulator^{3,4}, and show that memory retrieval is, indeed, partially context-independent. Moreover, we show that the mushroom bodies, which are required for olfactory^{5–7} but not visual or tactile learning⁸, effectively support context generalization. In visual learning in *Drosophila*, it appears that a facilitating effect of context cues for memory retrieval is the default state, whereas making recall context-independent requires additional processing.

We have developed a visual-learning paradigm in which a tethered fly can be conditioned by heat to prefer certain flight orientations and to avoid others (Fig. 1a, b)³. Learning is measured as the fly's lasting orientation preference when heat is permanently switched off. Conceptually, heat is the reinforcer (unconditioned stimulus) and those orientations of the panorama that are paired with heat during training are taken as the conditioned stimulus (CS⁺); those paired with the absence of heat are the CS[−] (ref. 9). In this description, as in most discussions of similar experiments, one intuitively assumes the conditioned stimulus to be the stimulus that is explicitly paired with the unconditioned stimulus. However, as the conditions during training and test are kept constant, the animal could also successfully learn with an unspecific conditioned stimulus comprising, in addition to the explicitly paired stimulus, all other stimuli that coincide with the unconditional stimulus. In our case, the CS⁺ and CS[−] would contain, in addition to the respective orientations of the panorama, the colour of the illumination, the stationary overhead pattern of the torque meter, the noise and odour in the laboratory, and so on. Obviously, in real life where significant events recur under changing circumstances, unspecific conditioned stimuli would be of little use.

To assess the importance and extent of context generalization in *Drosophila*, we systematically changed certain features of the experimental set-up (context) between training and test (Fig. 1c). The manipulated context is the illumination of the panorama. First, the illumination is changed between a monochromatic colour and white light or between blue and green. Without a switch, in broad-band monochromatic light, wild-type (control) flies remember the preferred orientation as well as in white light (Fig. 2a). If the illumination is switched after the second training between white light and a colour, memory retrieval is unaffected (Fig. 2b), indicating that the acquired pattern preferences are at least partially invariant to chromatic changes in the illumination. Curiously, even normal flies show no conditioned pattern preference if the

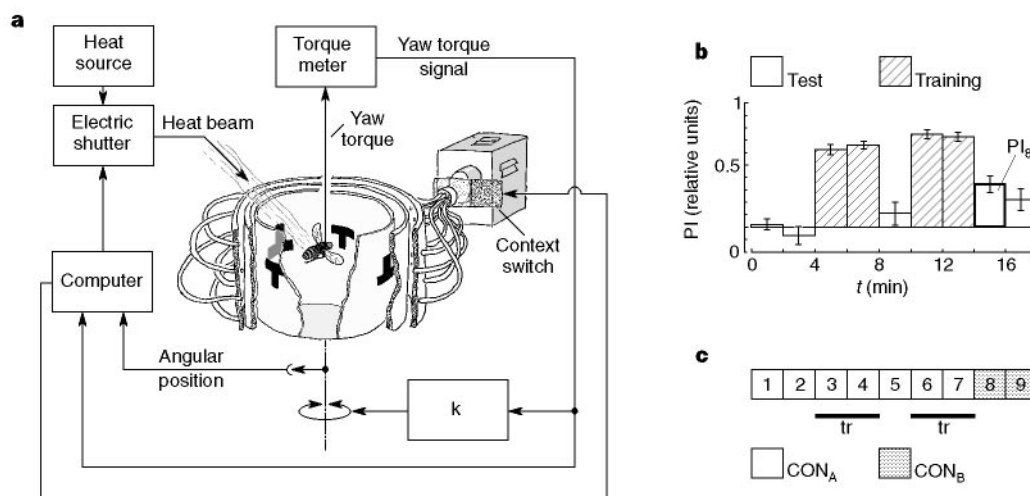


Figure 1 Experimental set-up and procedure. **a**, *Drosophila* flight simulator for testing visual learning with changing context (see Methods; box k calculates the angular velocity of the panorama from yaw torque). **b**, Performance indices (PI_{1-9}) of nine consecutive 2-min periods. Only the first 2-min period (PI_9) of the final learning test is shown in Figs 2–4. With few exceptions, PI_9 is smaller than PI_8

owing to the ongoing extinction training without heat. The PI_9 of this experiment is presented again in the first column of Fig. 2b. **c**, Time course of the experiment as used in subsequent experiments to indicate context conditions (CON_A and CON_B). The nine squares represent the recording periods and the horizontal bars underneath indicate training phases (tr).

illumination is switched between the two colours (Fig. 2c). A switch between two monochromatic colours may be a more severe context change than the switch between white light and colour, as the spectrum of the white light fully overlaps that of the respective colour.

In the experiment shown in Fig. 3 the illumination is white light, but it is turned off for 200 ms every 3 s (dark flashes). This regularly recurring disturbance has no detrimental effect on learning or short-term memory in the wild type (Fig. 3b). Moreover, switching from dark flashes to no dark flashes or vice versa between training and test does not interfere with memory retrieval (Fig. 3c, e). This and the previous examples show that memory retrieval in *Drosophila* is robust with respect to some but not all context changes.

To gain insight into the neural mechanisms of context generalization we tried to relate it to structural features of the brain. During the last 20 years, several routes for growing flies without functional mushroom bodies have been explored^{5–7,10,11}. As these different methods interfere with distinct developmental processes, a comparison of flies without mushroom bodies generated in different ways allows us to conclusively assign behavioural defects to the impairment of the mushroom bodies (as opposed to undesired side effects of a given method in other parts of the brain or body).

Flies without functional mushroom bodies perform well in many behaviours. They show coordinated walking, the full male courtship sequence, visual flight control and basic responses to olfactory and gustatory stimuli^{5–7,11}. They can perform many forms of learning such as visual pattern recognition, colour discrimination learning, motor learning, conditioned courtship suppression and spatial learning in the dark⁸. So far, only two behavioural properties have been shown to require mushroom bodies: odour discrimination learning^{5–7} and the control of spontaneous walking activity¹¹. How

the mushroom bodies support these behaviours is not yet known.

We have studied three types of mushroom body impairment. (1) In mutant *mushroom body miniature*¹ (*mbm*¹) females, the fibres of the intrinsic mushroom body neurons (Kenyon cells) degenerate during the third larval instar¹². (2) We ablated the mushroom body neuroblasts by applying the cytostatic drug hydroxyurea during the early first larval instar^{6,10}. (3) We targeted expression of the transgene for tetanus toxin light chain (*UAS_{Gal4}-CntE*)¹³ to a subset of Kenyon cells using the enhancer-trap line *P[Gal4]17D*¹¹.

Flies without mushroom bodies learn well in monochromatic blue or green light (Fig. 2a) as well as under dark-flash conditions (Fig. 3b). Thus, in normal and mushroom-body-ablated flies, learning is possible under a variety of experimental conditions as long as these conditions persist during training and test. To verify the absence of mushroom bodies, the mutant *mbm*¹ and hydroxyurea-treated flies used in the experiment were subjected to post mortem histology. In *P[Gal4]17D/UAS_{Gal4}-CntE* flies, toxin expression was monitored by immunohistochemistry^{11,13}. The latter flies could only be used for the first experiment (Fig. 2) as they did not show learning with dark flashes applied throughout the experiment (data not shown). The reason for this special sensitivity is not understood. It is probably due to toxin expression (outside the mushroom bodies) because the *P[Gal4]17D/CantonS* control flies did not show this behavioural deficit.

All three types of mushroom-body impairments led to a block of memory retrieval if any of the experimental conditions were changed between training and test. Flies tolerated neither a switch between monochromatic and white light (Fig. 2c) nor between dark flashes and no dark flashes (Fig. 3c, e). Considering that in complete

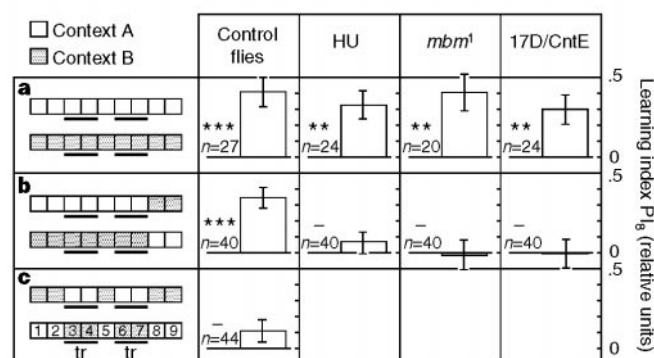


Figure 2 Influence of colour changes on memory retrieval. Three spectrally distinct illumination conditions (blue, green, white) are used in pairs (context A; context B). **a**, No context change between training and memory test. Under monochromatic blue and green light (data pooled) flies with and without mushroom bodies show high memory scores. **b**, Illumination is switched from monochromatic to white light or vice versa after the second training. Data of the four schedules (green to white; white to green; blue to white; white to blue) are pooled because none of them are statistically different either in pairs or alone. Control flies in this case are WTB treated exactly like hydroxyurea (HU)-treated flies except that HU is omitted from the procedure. For the *P[Gal4]17D/UAS_{Gal4}-CntE* flies, *P[Gal4]17D/CantonS* heterozygotes are used as controls. They show normal memory scores for all four conditions (data not shown). We did not use the mutated transgene *UAS_{Gal4}-ImpTNT* because, in a different set of experiments, it gave unexpected results. Note that although in *P[Gal4]17D/UAS_{Gal4}-CntE* flies tetanus toxin is detected only in Kenyon cells of the α/β type but not in α'/β' and γ fibres (for new nomenclature of mushroom body subsystems see ref. 29), the block in memory retrieval is as complete as in *mbm*¹ and HU flies. **c**, Between test and training phases illumination is switched from green to blue or vice versa (data pooled for both conditions). WTB flies show no memory retrieval.

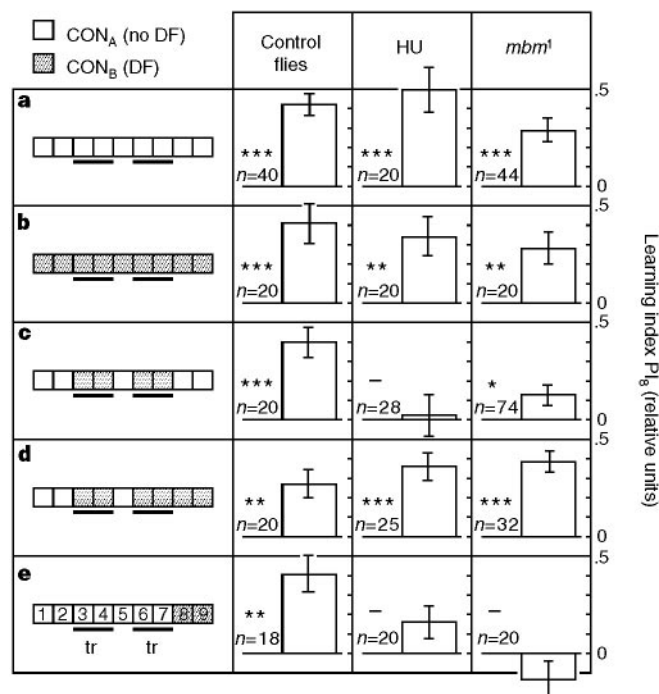


Figure 3 Context changes between dark flash (DF) condition and no dark flashes. **a**, Standard conditions. **b**, If DF condition is kept throughout the experiment it has no effect on memory retrieval in either controls or flies lacking mushroom bodies. Only *P[Gal4]17D/UAS_{Gal4}-CntE* flies fail to reveal any memory under continuous DF conditions (data not shown). **c**, If DFs are applied only during training, memory retrieval is severely impaired in flies lacking mushroom bodies but is normal in controls. **d**, Conditions as in **c** except that DFs continue after the second training. HU- and *mbm*¹ flies show no memory deficit. **e**, If DFs are applied only after the second training phase, the *Pi8* of flies without mushroom bodies is not significantly different from zero, whereas control flies still show normal memory retrieval.

darkness flies have no control of their orientation in the flight simulator, the latter result is particularly revealing. Omitting this severe interference with flight control after the training largely blocked memory retrieval, whereas continuing dark flashes had no effect (Fig. 3b, d). In Fig. 3c mutant *mbm*¹ flies, in contrast to hydroxyurea-treated flies, showed some memory traces after the dark flashes ceased. To test whether this effect was real, 74 *mbm*¹ flies were recorded under these conditions. Even if only flies with no or minute mushroom bodies were included in the evaluation, a small but significant memory score remained. This value, however, was significantly reduced compared with the experiments without context changes (difference between PI_8 in Fig. 3c and 3d; $P < 0.005$, Welch's unpaired *t*-test).

These results show that the mushroom bodies support memory retrieval after context changes. Under certain conditions we have observed delayed memory retrieval in flies without mushroom bodies. For instance, in hydroxyurea-treated flies, after a switch from no dark flashes to dark flashes, the last memory score (PI_9) was significantly positive ($PI_9 = 0.37 \pm 0.08$), supporting the idea that brain structures other than the mushroom bodies contribute to context generalization, albeit perhaps more slowly.

For efficient learning in a continuously changing world, not only the context but also the conditioned stimulus has to be generalized to some extent¹⁴. We investigated whether the mushroom bodies might also be required for this type of processing. A particular example of stimulus generalization has been described in visual pattern recognition¹⁵. In this experiment, composite patterns were used instead of the T-shaped patterns of Fig. 1. Between training and test the pattern elements were displaced up or down and adjusted in size to leave the vertical position of the centres of gravity of the patterns largely unchanged. Flies preferred or avoided the modified patterns according to the position of the centres of gravity of the patterns used in the training. It was shown in additional conditioning experiments (not shown) that the flies could discriminate the modifications introduced between training and test. In the experiment shown in Fig. 4, flies without mushroom bodies were subjected to this stimulus generalization task. They were clearly able to recognize the patterns according to the position of their centres of gravity, and their performance was not significantly different from that of control flies. We conclude that stimulus generalization in visual pattern recognition occurs independently of the mushroom bodies.

In current theories of memory retrieval, relying mainly on work in vertebrates, the separation of context and conditioned stimulus is taken largely for granted and only the facilitating role of context cues is emphasized^{1,2}. Formally, in *Drosophila* some context cues could be said to facilitate memory retrieval, as in vertebrates (blue versus green). However, in this experiment the change in colours between training and memory test determines whether the cue 'colour' of the preceding training has an effect on recall. Moreover,

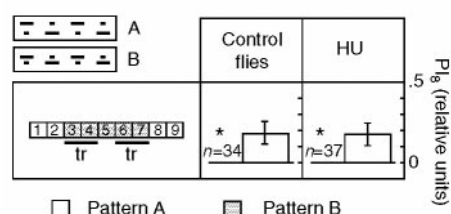
in the case of a change between blue and green, we can only observe the suppression, not the facilitation of recall. It seems more adequate, for this study, to treat the context effects on memory retrieval as a pattern-recognition problem.

Animals associate sensory signals with important events in order to recognize them immediately or even anticipate them once they recur. The simplest 'predictor' is a stored representation of the complete sensory pattern coinciding with or immediately preceding the event¹. Using such a stored representation (memory template) for recognition is a template-matching problem. If the memory template is very detailed and has to be precisely matched, the system will generate few false positives but will miss many of the events it was supposed to predict. If, on the other hand, the matching process is sloppy, it will identify many of the relevant events but also many false positives. To avoid this dilemma the animal needs to make the template as specific a predictor of the event as possible. In other words, it has to do what the learning psychologist does conceptually: separate the conditioned stimulus from the context. We have shown that, in visual learning at the flight simulator, the fly separates the conditioned stimuli (certain pattern orientations) from such context cues (colour of illumination, dark flashes and so on). We propose that it classifies as contextual, and therefore eliminates from the memory template, those stimuli which in their temporal structure bear no relation to the reinforcer. Although the importance of the temporal relation between the conditioned and unconditioned stimuli in classical conditioning is generally recognized⁹ (for *Drosophila*, see for instance ref. 16), little attention has been paid to the need for a more elaborate temporal analysis of the sensory stimuli as a means to distinguish the conditioned stimulus from the context.

So far, we cannot explain why a switch from one monochromatic light to another should block memory retrieval, whereas a switch between monochromatic and white light does not. Flies, like honeybees, have two receptor subsystems that are preferentially sensitive to blue and green light, respectively, and serve different visual functions¹⁷. It is possible that both subsystems independently mediate visual pattern recognition but have no crosstalk at the memory level. However, it is more likely that an intensely coloured environment is a salient cue and colour cues are only partially eliminated from the memory template. As studies in vertebrates indicate, complete context independence is maladaptive. An interesting further explanation would be that the flies can form an association between the colour and the possibility of being heated.

Our analysis shows that the mushroom bodies reduce context sensitivity and we have argued that this process requires the continuous temporal analysis of all the incoming sensory stimuli. The mushroom bodies are only two synapses away from the olfactory receptors and, in honeybees, cockroaches and crickets, they process multimodal sensory signals^{18–20}. On the basis of their special cytoarchitecture, mushroom bodies have been proposed to

Figure 4 Stimulus generalization occurs in flies without mushroom bodies. As an example of stimulus generalization in pattern vision, composite patterns of a long and a short horizontal bar (instead of T-shaped patterns)¹⁵ are used. In pattern set A the long bars are $30^\circ \times 10^\circ$ and the short bars $10^\circ \times 10^\circ$. In one pattern of a set the short bar is above, the long bar below the equator, in the other one this relation is inverted. The vertical separation of the bars is 30° . In pattern set B the long bars are $35^\circ \times 10^\circ$ and the short bars $8^\circ \times 10^\circ$. The vertical separation is 10° . In each set the two composite figures are discriminated by the different heights of their centres of gravity. With the switch from pattern set A to B between training and test the respective centres of gravity remain at about the same heights, whereas the pattern elements are displaced up and down. Both normal and mushroom body-less flies can still retrieve the learned pattern preferences after the switch, and retain their pattern preferences according to the heights of the centres of gravity of the patterns. (Control experiments show¹⁵ that the corresponding composite patterns of sets A and B can be discriminated.)



be involved in timing processes^{21,22}. However, alternative ways in which they might support memory retrieval in new contexts cannot be excluded. Also, how context generalization is related to the other functions described for the mushroom bodies^{5–7,11}, and how they operate as a network component in the brain remains to be understood. It has been suggested^{23,24} that the multimodal input from the protocerebrum to the mushroom body lobes serves a phylogenetically old function, whereas the salient chemosensory input from the antennal lobe to the mushroom body calyx may have been acquired secondarily. If this distinction holds, context generalization may well be an ancient mushroom body function and olfactory processing a more recent one.

No direct anatomical input from visual centres to the mushroom bodies has yet been discovered in *Drosophila* and, so far, the only functional link to the visual system has been the structural plasticity of the calyx induced by visual stimulation of the ipsilateral eye²⁵. Here, we have presented a behavioural account of visual processing in the mushroom bodies. In flies this modality offers many advantages for quantitative systems analysis¹⁷. The mushroom bodies may become the first central brain structure for which functional correlates at the behavioural level can be worked out in detail. □

Methods

Fly maintenance followed standard protocols as described²⁶. Wild-type 'Berlin' (WTB) and the mutant *mushroom body miniature*¹ (*mbm*¹) on WTB genetic background were used. In the present *mbm*¹ line more than 90% of the female flies have minute or no detectable mushroom bodies. The enhancer-trap line P[Gal4]17D was from a collection of H. Pfister and T. Raabe and has been described¹¹. The UAS_{Gal4}-CnTE line¹³ was provided by S. T. Sweeney and C. J. O'Kane. These two stocks had a wild-type CantonS genetic background. Two-to-six-day-old females were used in the experiments. For tethered flight, flies were briefly immobilized by cooling to 4 °C and were glued by their head and thorax to a small hook (copper wire; diameter 0.05 mm). After return to room temperature they were kept individually in small moist chambers overnight, with a few grains of sucrose. Immediately after the experiment, mutant and hydroxyurea-treated flies were anaesthetized by cooling, the hooks were ripped off with the glue and the flies were submerged in the fixative. We used standard paraffin histology for autofluorescence²⁷.

The apparatus for studying visual learning and context generalization is shown in Fig. 1a. The tethered fly is suspended at a torque meter in the centre of an arena that is illuminated from behind and carries four black patterns, for instance (Figs 1–3) two upright and two inverted T-shaped patterns in alternating sequence, on its wall. The Ts measure 36° vertically and 39° horizontally; the bars of the Ts are 12° wide. The light source is a 100W, 12V tungsten-iodine bulb. In the experiments of Figs 1 and 2 the following filters are used: (1) broad-band blue (Kodak Wratten gelatin filter No. 47); (2) broad-band green (Kodak Wratten gelatin filter No. 99); and (3) daylight filter BG18 (glass, 3 mm) from Schott, Mainz ('white').

The arena is rotated by an electric motor so that its angular velocity is proportional to but directed against, the fly's yaw torque (box k in Fig. 1); that is, the fly's yaw torque determines the angular velocity of the arena instead of rotating the fly's own body. This arrangement allows the tethered fly to stabilize and choose its flight orientation with respect to the arena by adjusting its yaw torque. A computer continuously registers the yaw torque and angular position of the arena. To test visual learning, a beam of infrared light is directed at the fly as an instantaneous source of heat. The beam can be intercepted by a computer-driven electric shutter. The arena is virtually divided into four quadrants with the patterns at their respective centres. During training the computer opens the shutter whenever the fly is heading into a quadrant with (for example) an upright T, and closes it when the fly is oriented towards one of the two adjacent quadrants. Hence, half of all possible orientations in the arena are paired with heat, the other with ambient temperature. During tests heat is permanently switched off. The 'context switch' controls the colour of the illumination or allows us to apply 'dark flashes' (see text).

Angular position is recorded every 50 ms, and orientation preferences are calculated for nine consecutive 2-min periods (performance index, PI_{1–9}).

Pattern A is paired with ambient temperature during training and pattern B with heat. Each of the two patterns (the upright and inverted T in the experiments of Figs 1–3) is pattern A in half of the experiments. This procedure eliminates the effects of spontaneous or non-associatively induced pattern preferences. If t_A is the time the fly spends heading towards the quadrants of pattern A, and t_B the time heading towards pattern B quadrants, the performance index (PI) is calculated as $PI = (t_A - t_B)/(t_A + t_B)$. Mean PIs are obtained for nine consecutive 2-min periods (PI_{1–9}). Before each 2-min period the arena is rotated to an arbitrary position. Further technical details regarding the flight simulator and data processing routines can be found in refs 7 and 27. Error bars in the figures are s.e.m.s, asterisks indicate levels of significance against zero. *** $P < 0.0005$; ** $P < 0.005$; * $P < 0.05$; – $P > 0.05$ (one-sample t -test; two-sided P -value). If pooled data are presented, none of the subgroups are statistically distinguishable.

To ablate mushroom body neuroblasts by hydroxyurea, we followed the procedure of ref. 10 as modified in ref. 8. Immunohistochemical methods were as described²⁸; the details for the tetanus toxin antibody are given in ref. 11.

Received 1 April; accepted 14 June 1999.

1. Tarpay, R. M. *Contemporary Learning Theory and Research* (McGraw-Hill, New York, 1997).
2. Bouton, M. E. Context, time, and memory retrieval in the interference paradigm of Pavlovian learning. *Psychol. Bull.* **114**, 80–99 (1993).
3. Wolf, R. & Heisenberg, M. Basic organization of operant behavior as revealed in *Drosophila* flight orientation. *J. Comp. Physiol. A* **169**, 699–705 (1991).
4. Heisenberg, M. & Wolf, R. Reafferent control of optomotor yaw torque in *Drosophila melanogaster*. *J. Comp. Physiol. A* **163**, 373–388 (1988).
5. Heisenberg, M., Borst, A., Wagner, S. & Byers, D. *Drosophila* mushroom body mutants are deficient in olfactory learning. *J. Neurogenet.* **2**, 1–30 (1985).
6. de Belle, J. S. & Heisenberg, M. Associative odor learning in *Drosophila* abolished by chemical ablation of mushroom bodies. *Science* **263**, 692–695 (1994).
7. Connolly, J. B. et al. Associative learning disrupted by impaired G(S) signalling in *Drosophila* mushroom bodies. *Science* **274**, 2104–2107 (1996).
8. Wolf, R. et al. *Drosophila* mushroom bodies are dispensable for visual, tactile, and motor learning. *Learn. Mem.* **5**, 166–178 (1998).
9. Pavlov, I. P. *Conditioned Reflexes* (trans. Anrep, G. V.) (Oxford Univ. Press, London, 1927).
10. Prokop, A. & Technau, G. M. Normal function of the mushroom body defect gene of *Drosophila* is required for the regulation of the number and proliferation of neuroblasts. *Dev. Biol.* **161**, 321–337 (1994).
11. Martin, J.-R., Ernst, R. & Heisenberg, M. Mushroom bodies suppress locomotor activity in *Drosophila melanogaster*. *Learn. Mem.* **5**, 179–191 (1998).
12. de Belle, J. S. & Heisenberg, M. Expression of *Drosophila* mushroom body mutations in alternative genetic backgrounds: A case study of the mushroom body miniature gene (*mbm*). *Proc. Natl Acad. Sci.* **93**, 9875–9880 (1996).
13. Sweeney, S. T., Broadie, K., Keane, J., Niemann, H. & O'Kane, C. J. Targeted expression of tetanus toxin light chain in *Drosophila* specifically eliminates synaptic transmission and causes behavioral defects. *Neuron* **14**, 341–351 (1995).
14. Bickel, W. K. & Etzel, B. C. The quantal nature of controlling stimulus–response relations as measured in tests of stimulus generalization. *J. Exp. Anal. Behav.* **44**, 245–270 (1985).
15. Ernst, R. & Heisenberg, M. The memory template in *Drosophila* pattern vision at the flight simulator. *Vis. Res.* (in the press).
16. Tully, T. & Quinn, W. G. Classical conditioning and retention in normal and mutant *Drosophila melanogaster*. *J. Comp. Physiol. A* **157**, 263–277 (1985).
17. Heisenberg, M. & Wolf, R. in *Studies of Brain Function* Vol. XII (ed. Braitenberg, V.) (Springer, Berlin, 1984).
18. Rybak, J. & Menzel, R. Integrative properties of the Pe1 neuron, a unique mushroom body output neuron. *Learn. Mem.* **5**, 133–145 (1998).
19. Li, Y.-S. & Strausfeld, N. J. Morphology and sensory modality of mushroom body efferent neurons in the brain of the cockroach, *Periplaneta americana*. *J. Comp. Neurol.* **387**, 631–650 (1997).
20. Schildberger, K. Multimodal interneurons in the cricket brain: properties of identified extrinsic mushroom body cells. *J. Comp. Physiol. A* **154**, 71–79 (1984).
21. Schürmann, F. Bemerkungen zur Funktion der *Corpora pedunculata* im Gehirn der Insekten aus morphologischer Sicht. *Exp. Brain Res.* **19**, 406–432 (1974).
22. Erbre, J., Homberg, U. & Gronenberg, W. in *Anthropoid Brain: Its Evolution, Development, Structure and Function* (ed. Gupta, P.) (Wiley, New York, 1987).
23. Strausfeld, N. J., Hansen, L., Li, Y., Gomez, R. S. & Ito, K. Evolution, discovery, and interpretations of arthropod mushroom bodies. *Learn. Mem.* **5**, 11–37 (1998).
24. Ito, K. et al. The organization of extrinsic neurons and their implications in the functional roles of the mushroom bodies in *Drosophila melanogaster* Meigen. *Learn. Mem.* **5**, 52–77 (1998).
25. Barth, M. & Heisenberg, M. Vision affects mushroom bodies and central complex in *Drosophila melanogaster*. *Learn. Mem.* **4**, 219–229 (1997).
26. Guo, A. et al. Conditioned visual flight orientation in *Drosophila*: Dependence on age, practice and diet. *Learn. Mem.* **3**, 49–59 (1996).
27. Ashburner, M. *Drosophila. A Laboratory Manual. Protocol* 112. 254 (Cold Spring Harbor Lab. Press, Cold Spring Harbor, NY, 1989).
28. Buchner, E. et al. Cell-specific immuno-probes for the brain of normal and mutant *Drosophila melanogaster*. I. Wild type visual system. *Cell Tissue Res.* **253**, 357–370 (1988).
29. Crittenden, J. R., Skoulakis, E. M. C., Han, K.-A., Kalderon, D. & Davis, R. L. tripartite mushroom body architecture revealed by antigenic markers. *Learn. Mem.* **5**, 38–51 (1998).

Acknowledgements. We thank S. Kramer for critically reading the manuscript; S. Clemens-Richter for technical help; T. Raabe for the enhancer-trap line P[Gal4]17D; S. T. Sweeney and C. J. O'Kane for the UAS_{Gal4}-CnTE stock; and J. Thierier and H. Niemann for tetanus toxin antibody. The work has been supported by Deutsche Forschungsgemeinschaft, Fond der Chemischen Industrie (M.H.) and Alexander von Humboldt-Foundation (L.L.).

Correspondence and requests for materials should be addressed to M.H. (e-mail: heisenberg@biozentrum.uni-wuerzburg.de).

Synaptic function modulated by changes in the ratio of synaptotagmin I and IV

J. Troy Littleton^{*†}, Thomas L. Serano^{†‡}, Gerald M. Rubin[‡], Barry Ganetzky^{*} & Edwin R. Chapman[§]

^{*} Laboratory of Genetics and [§] Department of Physiology, University of Wisconsin, Madison, Wisconsin 53706, USA

[‡] Department of Molecular and Cell Biology, Howard Hughes Medical Institute, University of California, Berkeley, California 94720-3200, USA

[†] These authors contributed equally to this work.

Communication within the nervous system is mediated by Ca^{2+} -triggered fusion of synaptic vesicles with the presynaptic plasma membrane. Genetic and biochemical evidence indicates that synaptotagmin I may function as a Ca^{2+} sensor in neuronal exocytosis because it can bind Ca^{2+} and penetrate into lipid bilayers^{1–4}. Chronic depolarization or seizure activity results in the upregulation of a distinct and unusual isoform of the synaptotagmin family, synaptotagmin IV (ref. 5). We have identified a *Drosophila* homologue of synaptotagmin IV that is enriched on synaptic vesicles and contains an evolutionarily conserved substitution of aspartate to serine that abolishes its ability to bind membranes in response to Ca^{2+} influx. Synaptotagmin IV forms hetero-oligomers with synaptotagmin I, resulting in synaptotagmin clusters that cannot effectively penetrate lipid bilayers and are less efficient at coupling Ca^{2+} to secretion *in vivo*: upregulation of synaptotagmin IV, but not synaptotagmin I, decreases evoked neurotransmission. These findings indicate that modulating the expression of synaptotagmins with different Ca^{2+} -binding affinities can lead to heteromultimers that can regulate the efficiency of excitation–secretion coupling *in vivo* and represent a new molecular mechanism for synaptic plasticity.

Synaptotagmins are synaptic vesicle proteins of which 12 isoforms have been identified in mammals, including many with different Ca^{2+} -binding properties⁶. Here we report three new synaptotagmins from *Drosophila* and *Caenorhabditis elegans* that are homologues of mammalian synaptotagmins IV, V and VII. Synaptotagmins contain two cytoplasmic repeats homologous to the C2 domains found in Ca^{2+} -dependent isoforms of protein kinase C⁷. The C2A domain binds Ca^{2+} and anionic lipids, and triggers the penetration of synaptotagmin I into membranes^{4,8}, whereas the C2B domain mediates Ca^{2+} -triggered oligomerization⁹. These C2 domains cooperate to form high-affinity Ca^{2+} -dependent complexes with the plasma membrane t-SNAREs syntaxin and SNAP-25 (refs 9–13).

Administration of kainic acid in rats results in clinical features of epilepsy and status epilepticus, and causes a selective and dramatic increase in an unusual isoform of synaptotagmin, synaptotagmin IV (ref. 5). Both the *Drosophila* and *C. elegans* synaptotagmin IV homologue contain a conserved substitution of aspartate to serine in the third Ca^{2+} ligand in C2A¹⁴ (Fig. 1a). To determine whether *Drosophila* synaptotagmin IV can penetrate membranes in the presence of Ca^{2+} , we tested the ability of immobilized recombinant cytoplasmic domains of synaptotagmins I and IV to bind to anionic liposomes in the presence and absence of Ca^{2+} (Fig. 1b). Fusion proteins of *Drosophila* synaptotagmin I containing C2A and C2AB showed robust Ca^{2+} -dependent phospholipid binding, whereas fusion proteins of synaptotagmin IV containing C2A or C2AB failed to bind lipids. However, like synaptotagmin I, synaptotagmin IV also interacted with the clathrin adapter AP-2, the synprint domain from N-type Ca^{2+} channels, and the t-SNARE syntaxin

(Fig. 1d). The substitution of aspartate to serine is also found in mammalian synaptotagmins IV and XI (ref. 15) and phylogenetic analyses indicate that these synaptotagmins make up a subfamily of evolutionarily conserved synaptotagmins that fail to bind anionic lipids in response to Ca^{2+} (Fig. 1c).

Whole-mount embryonic *in situ* hybridization indicates that synaptotagmin IV and synaptotagmin I are expressed ubiquitously and specifically throughout the nervous system and are coexpressed in most, if not all, neurons (Fig. 1e). In order to investigate the subcellular distribution of synaptotagmin IV, we generated antisera against bacterially expressed recombinant synaptotagmin IV. The antisera recognize a protein of relative molecular mass 55,000 that is expressed in *Drosophila* head extracts but is not detected by preimmune sera. Synaptotagmin IV, like synaptotagmin I and synaptobrevin, is enriched in *Drosophila* synaptic vesicles, as assayed by subcellular fractionation of *Drosophila* heads (Fig. 2c). Synaptotagmin IV is also present on synaptic vesicles immunisolated using anti-cysteine-string protein antibodies (data not shown), and co-sediments with other *Drosophila* synaptic vesicle proteins on sucrose gradients (Fig. 2d). A shift in both synaptotagmin I and IV from synaptic vesicles to presynaptic membranes in *shibire* mutants at the non-permissive temperature confirms that both synaptotagmins are present on recycling synaptic vesicles and accumulate in the plasma membrane during a block of endocytosis (Fig. 2d). Colocalization of synaptotagmin I and IV on the same population of synaptic vesicles was confirmed by immunisolating synaptic vesicles from *Drosophila* head homogenates by using anti-synaptotagmin I antibodies and immunoblotting for synaptotagmin IV (Fig. 3f).

All the Ca^{2+} ligands are conserved in the C2B domain of *Drosophila* synaptotagmin IV (Fig. 1a), indicating that Ca^{2+} -triggered oligomerization may remain intact. As shown by Coomassie staining and immunoblotting, recombinant synaptotagmins are able to form hetero-oligomeric complexes *in vitro* (Fig. 3a, b). To test whether native synaptotagmin was also capable of forming hetero-oligomers, we prepared *Drosophila* head membranes and incubated them with either synaptotagmin I, IV or VII immobilized recombinant proteins. Native synaptotagmin I showed Ca^{2+} -dependent binding to recombinant synaptotagmins I, IV and VII, indicating that Ca^{2+} -dependent oligomerization may be a conserved property of the synaptotagmin family (Fig. 3c). Co-immunoprecipitation of synaptotagmin IV with anti-synaptotagmin I antibodies from solubilized *Drosophila* head membranes confirmed that there is a specific interaction between the two native proteins (Fig. 3d). We next determined how the assembly of synaptotagmins into I–I and I–IV oligomeric complexes affected their ability to interact with membranes in the presence of Ca^{2+} . Assembly of I–I oligomers resulted in a 45% increase in Ca^{2+} -dependent membrane binding (Fig. 3e), whereas assembly of I–IV oligomers decreased membrane binding by 35% compared with synaptotagmin I alone, and by 55% compared with I–I oligomeric complexes ($P > 0.05$, Student's *t*-test). Thus, the formation of synaptotagmin I–IV hetero-oligomers decreases the ability of synaptotagmin I to penetrate membranes, indicating that changes in the relative abundance of distinct synaptotagmin isoforms on the same synaptic vesicle may modulate the Ca^{2+} sensitivity of vesicle fusion *in vivo*.

To test directly the *in vivo* consequences of synaptotagmin IV upregulation on synaptic function, we used *elav*–GAL4 to drive UAS–synaptotagmin IV in a pan-neuronal manner in transgenic *Drosophila*. As a control, UAS–synaptotagmin I transgenic *Drosophila* were generated and analysed in an identical manner. Immunoblots demonstrated a fivefold overexpression of synaptotagmin I or synaptotagmin IV in fly strains harbouring *elav*–GAL4 (Fig. 4d). Electrophysiological recordings at the third-instar neuromuscular junction were performed in high-concentration extracellular Ca^{2+} (3 mM) for maximal stimulation of the release machinery. The peak amplitude of evoked responses from nerve stimulation was

decreased by 21% in UAS-*synaptotagmin IV*; *elav*-GAL4 larvae compared with controls or UAS-*synaptotagmin I*; *elav*-GAL4 larvae ($P > 0.05$, Student's *t*-test) (Fig. 4a). This decrease in evoked release was of presynaptic origin, as neither muscle miniature excitatory junction potential (MEJP) amplitude or muscle resting membrane potential was affected. Endocytosis and vesicle recycling also remained intact as high-frequency stimulation did not result in further decreases in evoked responses over time

(Fig. 4b). Whereas overexpression of synaptotagmin IV decreased evoked secretion, overexpression of either synaptotagmin I or IV resulted in decreases in MEJP frequency (Fig. 4c). Thus, both isoforms exhibit inhibitory 'clamp' actions on spontaneous fusion events, with isoforms I and IV playing active and inhibitory roles, respectively, in evoked secretion.

Our data indicate that synaptotagmin IV can downregulate the level of synaptic transmission owing to loss of Ca^{2+} -dependent

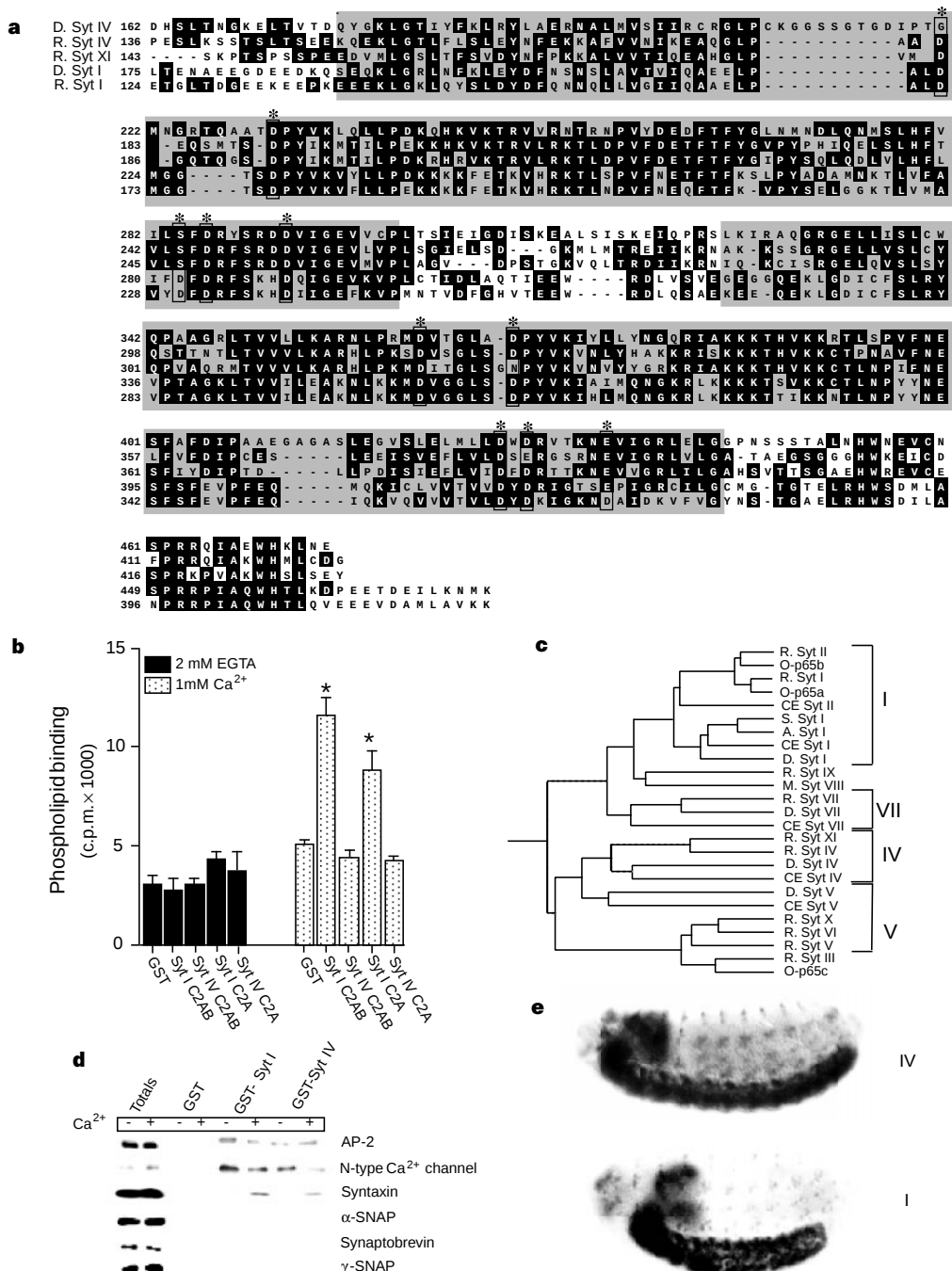


Figure 1 Identification of the *Drosophila* synaptotagmin family. **a**, Alignment of *Drosophila* synaptotagmins. Identities are shown in black, the two C2 domains in grey, and the five acidic Ca^{2+} ligands are boxed. Syt, synaptotagmin. **b**, Ca^{2+} -dependent phospholipid binding to immobilized recombinant synaptotagmin I or IV. **c**, Dendrogram of the synaptotagmin family. D, *Drosophila*; R, rat; O, *Discopyge ommata*; C, *C. elegans*; M, mouse; A, *Aplysia*; S, longfin squid. **d**, Binding of recombinant syntaxin and synprint peptides and native AP-2 μ -adaptin, α -SNAP

and γ -SNAP to recombinant *Drosophila* synaptotagmins. Synaptotagmin I and IV also bind inositol polyphosphates¹⁹. Interactions with SV2 have been reported for synaptotagmin I, but this interaction is disrupted by physiological concentrations of Mg^{2+} , and no SV2 homologue has yet been identified in *Drosophila*. **e**, Whole-mount *in situ* hybridization to stage 16 (synaptotagmin IV) or 17 (synaptotagmin I) *Drosophila* embryos with digoxigenin-labelled probes.

phospholipid binding activity, coupled with its ability to assemble into oligomers with synaptotagmin I, thereby allowing the synapse to adapt a new state of activity. Upregulation of synaptotagmin IV in response to seizure activity may have evolved as an adaptive mechanism to decrease neural activity. Although we cannot exclude the possibility that increased synaptotagmin IV expression may downregulate release by an unknown mechanism, the only biochemical difference we found between synaptotagmin I and IV is their ability to bind phospholipids in a Ca^{2+} -dependent manner. The large number of different synaptotagmin isoforms with unique Ca^{2+} -binding affinities, and the finding that mammalian synap-

tagmin isoforms can also assemble into heteromultimers^{16,17}, provide the potential for a large combinatorial 'library' of Ca^{2+} sensors that assemble *in vivo* to modulate synaptic strength. □

Methods

Binding assays and production of recombinant proteins and antibodies.

Complementary DNA encoding the cytoplasmic domains of *Drosophila* synaptotagmins I (C2A, residues 144–317; C2AB, residues 144–474), IV (C2A, residues 132–323; C2AB, residues 132–474) and VII (C2AB, residues 28–417) were expressed as glutathione S-transferase fusion proteins as described¹³. Thrombin was used to cleave synaptotagmin IV from GST and

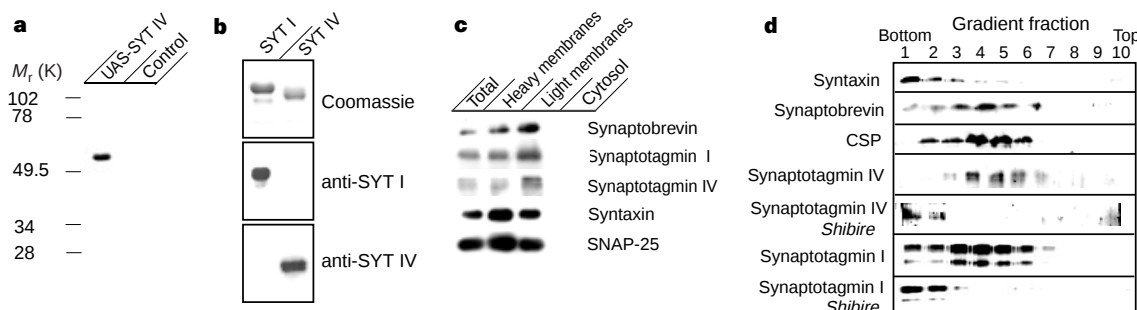


Figure 2 Synaptotagmin IV is a synaptic vesicle protein. **a**, Western blot of dissected imaginal discs from control or UAS-synaptotagmin IV; *sevenless*-GAL4 larva. The synaptotagmin IV antisera specifically recognize a protein of relative molecular mass 55,000 (M_r , 55K) from the *synaptotagmin IV* gene under the control of the UAS promoter. **b**, Top, Coomassie gel of synaptotagmins I and IV. Bottom, immunoblots probed with isoform-specific antibodies. **c**, Synaptotagmin IV is concentrated in light membrane fractions enriched in synaptic

vesicles. **d**, Sucrose gradient velocity sedimentation fractions of *Drosophila* head membranes from CS or heat-shocked *shibire* adult flies were probed for antigens as shown. Plasma membrane proteins such as syntaxin and the sodium pump migrate to the bottom of the gradient in the last two fractions. In contrast, synaptic vesicle proteins migrate in fractions 3–6. In *shibire* mutants, synaptic vesicle proteins accumulate in the plasma membrane.

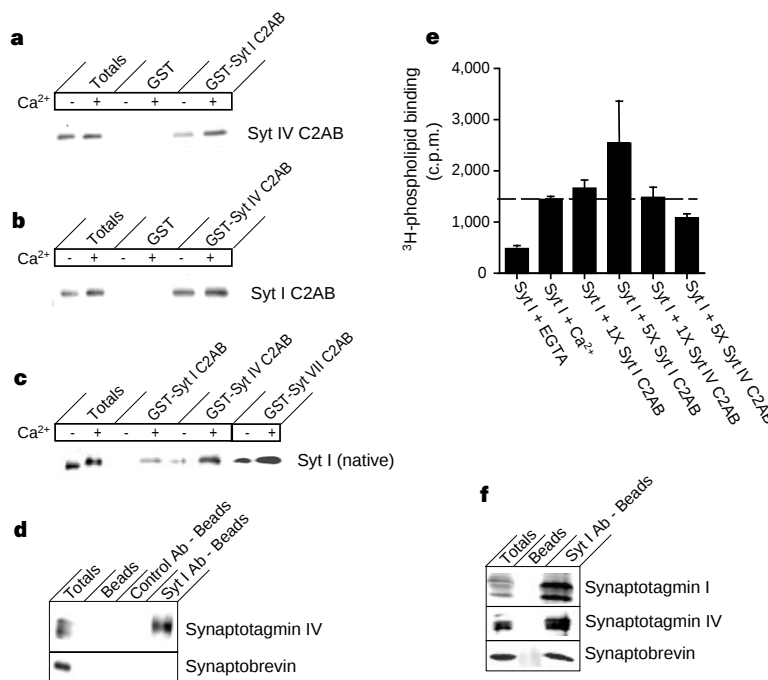


Figure 3 Hetero-oligomerization of *Drosophila* synaptotagmins I and IV decreases membrane binding. Binding was detected by Coomassie staining (**a**, **b**) or by immunoblotting using an anti-synaptotagmin I luminal domain antibody (**c**). **a**, Binding of 1 μM recombinant synaptotagmin (Syt) IV cytoplasmic domain (C2AB) to 30 μg of immobilized recombinant synaptotagmin I C2AB in 100 μl . **b**, Binding of 1 μM recombinant synaptotagmin I cytoplasmic domain to 30 μg immobilized recombinant synaptotagmin IV C2AB in 100 μl . **c**, Binding of native synaptotagmin I from *Drosophila* head detergent extracts to 30 μg

recombinant immobilized synaptotagmins I, IV and VII. **d**, Co-immunoprecipitation of synaptotagmin IV with synaptotagmin I from Triton X-100 solubilized *Drosophila* head membranes incubated in 1 mM Ca^{2+} using anti-synaptotagmin I antibodies cross-linked to protein G beads. **e**, Phospholipid binding to synaptotagmin I-IV hetero-oligomers. **f**, Synaptotagmin IV is found on synaptotagmin I immunoprecipitated vesicles. Similar results are seen with mammalian synaptotagmin IV (refs 17, 20).

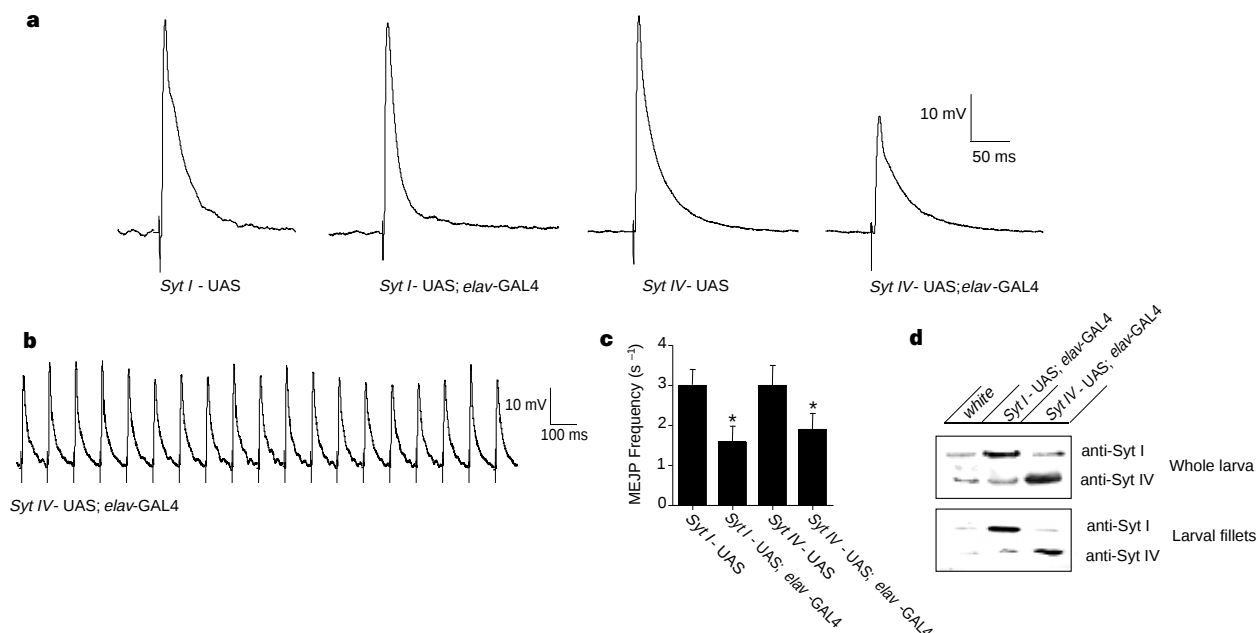


Figure 4 Electrophysiological analysis of transgenic *Drosophila* overexpressing synaptotagmin IV or synaptotagmin I. **a**, Transgenic larvae containing either *synaptotagmin I*-UAS or *synaptotagmin IV*-UAS alone or with *elav*-GAL4 to drive neuronal overexpression were recorded in 3 mM Ca²⁺ HL3 solution. Peak amplitude of the excitatory junctional potential and the resting membrane potential from muscle fibre 6 were as follows for each genotype: *synaptotagmin I*-UAS (39.3 ± 2.1, 56 ± 1.8, *n* = 8); *synaptotagmin I*-UAS; *elav*-GAL4 (40.6 ± 2.1, 60 ± 1.9, *n* = 17); *synaptotagmin IV*-UAS (41.4 ± 2.1, 60 ± 2, *n* = 14); *synaptotagmin IV*-UAS; *elav*-GAL4 (32.9 ± 2.6, 58 ± 1.7, *n* = 19). **b**, Stimulation

(10 Hz) of a neuromuscular junction from *synaptotagmin IV*-UAS; *elav*-GAL4 larvae showed no depletion over time. **c**, MEJP frequency was decreased by overexpression of either synaptotagmin I or IV (*P* < 0.05, Student's *t*-test). **d**, Five larvae of the indicated genotypes were homogenized in sample buffer. Immunoblots with anti-synaptotagmin I or anti-synaptotagmin IV antisera showed a specific fivefold upregulation of the corresponding protein in *elav*-GAL4 larva containing the UAS transgenes compared to controls. Similar levels of overexpression were observed in immunoblots of larval neuromuscular junction fillets, as used for electrophysiology (bottom).

the cleaved material was separated by SDS-PAGE and used to immunize rabbits according to standard procedures, using the University of Wisconsin polyclonal antibody facility. For testing binding of native synaptotagmin I to recombinant synaptotagmins, detergent extracts from *Drosophila* head membranes were prepared¹³ and 2 mg of extract was incubated with 30 µg recombinant protein, and 7.5% of total and 17% of bound material was loaded onto gels. Synaptotagmin homo- and hetero-oligomers were assembled by adding either equimolar or a fivefold molar excess of recombinant synaptotagmin I or IV to 15 µg immobilized synaptotagmin I for 2 h in 1 mM Ca²⁺. Unbound soluble synaptotagmin was removed by washing three times before samples were assayed for phospholipid binding activity. Velocity sedimentation was carried out using 10–30% sucrose gradients centrifuged for 1 h at 50 r.p.m. in a VTi65 rotor. Immunoprecipitation of synaptic vesicles and immunoprecipitation from *Drosophila* head membranes was as described¹⁸.

Generation of transgenic *Drosophila*. A 2-kilobase (kb) fragment encompassing the *synaptotagmin I* open reading frame (ORF) and a 3.1-kb fragment spanning the *synaptotagmin IV* ORF (obtained from the *Drosophila* Berkeley Genome Project) were subcloned into PUASt vectors, and transgenic *Drosophila* were generated by germline transformation.

Electrophysiological analysis. Electrophysiological recordings were performed on male third-instar non-*Tb* larva from crosses of UAS-*synaptotagmin I* or IV virgins to *elav*-GAL4; T(2:3) CyO-*TM6Tb* males in 3 mM Ca²⁺-HL3 as described² using an Axoclamp 2B.

Phospholipid binding assays. Phospholipid assays were performed with 25%PS/75%PC using 15 µg recombinant protein as described⁴.

Received 28 May; accepted 30 June 1999.

- Littleton, J. T., Stern, M., Schulze, K., Perin, M. & Bellen, H. J. Mutational analysis of *Drosophila synaptotagmin* demonstrates its essential role in Ca²⁺-activated neurotransmitter release. *Cell* **74**, 1125–1134 (1993).
- Littleton, J. T., Stern, M., Perin, M. & Bellen, H. J. Calcium dependence of neurotransmitter release and rate of spontaneous vesicle fusions are altered in *Drosophila synaptotagmin* mutants. *Proc. Natl Acad. Sci. USA* **91**, 10888–10892 (1994).
- Geppert, M. et al. Synaptotagmin I: a major Ca²⁺ sensor for transmitter release at a central synapse. *Cell* **79**, 717–727 (1994).

- Chapman, E. & Davis, A. Direct interaction of a Ca²⁺-binding loop of synaptotagmin with lipid bilayers. *J. Biol. Chem.* **273**, 13995–14001 (1998).
- Vician, L. et al. Synaptotagmin IV is an immediate early gene induced by depolarization in PC12 cells and in brain. *Proc. Natl Acad. Sci. USA* **92**, 2164–2168 (1995).
- Sudhof, T. & Rizo, J. Synaptotagmins: C2-domain proteins that regulate membrane traffic. *Neuron* **17**, 379–388 (1996).
- Perin, M. S., Fried, V. A., Mignery, G. A., Jahn, R. & Südhof, T. C. Phospholipid binding by a synaptic vesicle protein homologous to the regulatory region of protein kinase C. *Nature* **345**, 260–263 (1990).
- Brose, N., Petrenko, A. G., Südhof, T. C. & Jahn, R. Synaptotagmin: a Ca²⁺ sensor on the synaptic vesicle surface. *Science* **256**, 1021–1025 (1992).
- Chapman, E. R., An, S., Edwardson, J. M. & Jahn, R. A novel function for the second C2 domain of synaptotagmin: Ca²⁺-triggered dimerization. *J. Biol. Chem.* **271**, 5844–5849 (1996).
- Chapman, E. R., Hanson, P. L., An, S. & Jahn, R. Ca²⁺ regulates the interaction between synaptotagmin and syntaxin 1. *J. Biol. Chem.* **270**, 23667–23671 (1995).
- Schiavo, G., Stenbeck, G., Rothman, J. & Sollner, T. Binding of the synaptic vesicle v-SNARE, synaptotagmin, to the plasma membrane t-SNARE, SNAP-25, can explain docked vesicles at neurotoxin-treated synapses. *Proc. Natl Acad. Sci. USA* **94**, 997–1001 (1997).
- Li, C. et al. Ca²⁺-dependent and -independent activities of neural and non-neural synaptotagmins. *Nature* **375**, 594–599 (1995).
- Littleton, J. T. et al. Temperature-sensitive paralytic mutations demonstrate that synaptic exocytosis requires SNARE complex assembly and disassembly. *Neuron* **21**, 401–413 (1998).
- Sutton, R., Davletov, B., Berghuis, A., Südhof, T. & Sprang, S. Structure of the first C2 domain of synaptotagmin I: a novel Ca²⁺/phospholipid-binding fold. *Cell* **80**, 929–938 (1995).
- von Poser, C., Ichchenko, K., Shao, X., Rizo, J. & Südhof, T. The evolutionary pressure to inactivate: A subclass of synaptotagmins with an amino acid substitution that abolishes Ca²⁺ binding. *J. Biol. Chem.* **272**, 14314–14319 (1997).
- Chapman, E., Desai, R., Davis, A. & Tornehl, C. Delineation of the oligomerization, AP-2 binding, and synprint binding region of the C2B domain of synaptotagmin. *J. Biol. Chem.* **273**, 32966–32973 (1998).
- Osborne, S. L., Herreros, J., Bastiaens, P. I. & Schiavo, G. Calcium-dependent oligomerization of synaptotagmins I and II. *J. Biol. Chem.* **274**, 59–66 (1999).
- van de Goor, J., Ramaswami, M. & Kelly, R. Redistribution of synaptic vesicles and their proteins in temperature-sensitive shibire(ts1) mutant *Drosophila*. *Proc. Natl Acad. Sci. USA* **92**, 5739–5743 (1995).
- Fukuda, M., Kojima, T., Aruga, J., Niinobe, M. & Mikoshiba, K. Functional diversity of C2 domains of synaptotagmin family. *J. Biol. Chem.* **270**, 26523–26527 (1995).
- Ferguson, G. D., Thomas, D. M., Elferink, L. A. & Herschman, H. R. Synthesis, degradation, and subcellular localization of synaptotagmin IV, a neuronal immediate early gene product. *J. Neurochem.* **72**, 1821–1831 (1999).

Acknowledgements. We thank R. Kelly, S. Benzer, L. Brodin, C. Goodman, W. Catterall, H. Jackle and R. Jahn for antibodies and reagents, and J. Pendleton for help with sequencing. This work was supported by grants from the NIH and the Howard Hughes Medical Institute. T.L.S. is a Jane Coffin Childs postdoctoral fellow. J.T.L. is a Merck fellow of the Helen Hay Whitney Foundation.

Correspondence and requests for materials should be addressed to J.T.L. (e-mail: jtittle@facstaff.wisc.edu).

Interaction of glutamic-acid-rich proteins with the cGMP signalling pathway in rod photoreceptors

Heinz G. Körschen*, Michael Beyermann†, Frank Müller*, Martin Heck‡, Marius Vantler*, Karl-Wilhelm Koch*, Roland Kellner§, Uwe Wolftrum||, Christian Bode||, Klaus Peter Hofmann‡ & U. Benjamin Kaupp*

* Institut für Biologische Informationsverarbeitung, Forschungszentrum Jülich, 52425 Jülich, Germany

† Forschungsinstitut für Molekulare Pharmakologie, 10315 Berlin, Germany

‡ Institut für Medizinische Physik und Biophysik, Universitätsklinikum Charité, Humboldt-Universität, 10998 Berlin, Germany

§ Institut für Physiologische Chemie und Pathobiochemie, Universität Mainz, 55099 Mainz, Germany

|| Institut für Zoologie, Universität Karlsruhe, 76128 Karlsruhe, Germany

The assembly of signalling molecules into macromolecular complexes (transducinomes) provides specificity, sensitivity and speed in intracellular signalling pathways^{1,2}. Rod photoreceptors in the eye contain an unusual set of glutamic-acid-rich proteins (GARPs) of unknown function^{3–7}. GARPs exist as two soluble forms, GARP1 and GARP2, and as a large cytoplasmic domain (GARP' part) of the β -subunit of the cyclic GMP-gated channel^{3–7}. Here we identify GARPs as multivalent proteins that interact with the key players of cGMP signalling, phosphodiesterase and guanylate cyclase, and with a retina-specific ATP-binding cassette transporter (ABCR)^{8,9}, through four, short, repetitive sequences. In electron micrographs, GARPs are restricted to the rim region and incisures of discs in close proximity to the guanylate cyclase and ABCR, whereas the phosphodiesterase is randomly distributed. GARP2, the most abundant splice form, associates more strongly with light-activated than with inactive phosphodiesterase, and GARP2 potently inhibits phosphodiesterase activity. Thus, the GARPs organize a dynamic protein complex near the disc rim that may control cGMP turnover and possibly other light-dependent processes. Because there are no similar GARPs in cones, we propose that GARPs may prevent unnecessary cGMP turnover during daylight, when rods are held in saturation by the relatively high light levels.

The β -subunit of the cGMP-gated channel has a unique bipartite structure⁴; it embodies a membrane-spanning region (β' part) and a large cytosolic amino-terminal region (GARP' part) (Fig. 1a). The β' part contains all of the functional motifs characteristic of cGMP-gated channel subunits, whereas the GARP' part is largely identical with a soluble splice form, GARP1 (refs 3, 4). Another splice form, GARP2, lacks the carboxy-terminal glutamic-acid-rich region^{5,6} (Fig. 1a).

We found that antibodies, each specific for one GARP form (see epitopes in Fig. 1a), recognized polypeptides with a relative molecular mass (M_r) of ~240,000 (240K; channel β), 130K (GARP1) and 62K (GARP2), respectively, in western blots of rod outer segments (ROS) (Fig. 1b). The differences between the calculated and apparent M_r s of GARPs result from the anomalous electrophoretic mobility caused by the large number of Glu residues, which reduce SDS binding⁴. A comparison of the staining intensity produced by a common anti-GARP antibody suggests that GARP2 is much more abundant than GARP1 (Fig. 1b; see also ref. 6). Moreover, by

comparing the staining intensity of the 62K band with that of known amounts of recombinant GARP2 (rGARP2), we estimate that GARP2 is at least as abundant as phosphodiesterase (Rh: GARP2 $\approx$ 100 : 1; data not shown). Thus, GARP2 is a major protein of ROS. We separated ROS proteins into soluble and membrane fractions, and tested for the presence of GARPs by western blotting. Hypotonic washing resulted in the loss of GARP1 and GARP2 from the membrane fraction and complete recovery in the soluble fraction (Fig. 1c; compare lanes HMF and HSF). In contrast, isotonic washing did not remove GARPs from membranes (compare lanes IMF and ISF). Under isotonic conditions, soluble GARP1 and GARP2 completely reassociate with membranes depleted of both proteins (Fig. 1c; compare lanes RMF and RSF). Thus, like transducin and phosphodiesterase, GARP1 and GARP2 are tightly bound to the membrane at physiological conditions.

In vertical cryosections of bovine retina (Fig. 1d), anti-GARP stained outer segments and spherules of rods (Fig. 1d, arrowheads), but not other retinal cells, including cone photoreceptors. This finding was corroborated using freshly dissociated rods and cones. Whereas the outer segment and spherule of solitary rods were labelled (Fig. 1e), we detected no staining in any compartment of solitary cones (Fig. 1f, arrowhead). This indicates that GARPs may have a function that is specific to rods and absent in cones, although it is possible that cones express distantly related GARPs not recognized by the antibodies.

The common N-terminal region of GARPs contains four consecutive proline-rich repeats, R1–R4 (Fig. 1g). The repeat motif is characterized by an invariant Trp residue and a Pro–Gln–Pro triplet separated by nine, mostly conserved, residues, except that the first repeat is shorter than the others. Because R1–R4 are the most conserved sequence elements among GARPs from different species^{3–7}, we reasoned that these repeats serve as targets for binding other proteins. To identify potential target proteins, we constructed affinity columns from peptides, each representing one of the four repeats. ROS proteins were separated into cytosolic and membrane fractions and tested for their ability to bind to the peptide columns (Fig. 2). Phosphodiesterase (α , β ; 90K) is the most abundant cytosolic protein retained by the columns, then eluted by SDS buffer (Fig. 2a). The 37K α -subunit (T_α) (Fig. 2a) and the 35K β -subunit (T_β) (data not shown) of transducin bound weakly to repeat columns, while GARP1 (130K) and GARP2 (62K) bound quantitatively (Fig. 2a). Polypeptide binding to R1–R4 was specific; binding to an unrelated peptide or an empty column was absent or insignificant (Fig. 2a, lanes C1, C2). Binding was also specific inasmuch as other proteins, for example protein kinase C (Fig. 2a), recoverin or arrestin (data not shown), did not bind to the columns.

When we used stripped, detergent-solubilized ROS membranes, polypeptides of M_r s 63K, 112K, 220K and 240K bound to the columns. The 63K and 240K polypeptides have been identified as the α - and β -subunit, respectively, of the rod cGMP-gated channel (Fig. 2b), the 112K polypeptide as guanylate cyclase (Fig. 2b) and the 220K polypeptide as ABCR^{8–10}. Rhodopsin, the most abundant protein, did not specifically bind to the R1–R4 peptides. The repeats are not equal in their potency in binding target proteins. For example, R2 and R3 bind phosphodiesterase strongly, whereas R1 and phosphodiesterase interact weakly, if at all (Fig. 2a). However, we found no consistent pattern of target preference to suggest that each repeat recruits a specific protein, as has been found for PDZ domains in the InaD scaffolding protein of *Drosophila* photoreceptors¹¹. Thus, more than one repeat or the tetrad repeat unit might be involved in binding target proteins.

Using purified components, we examined the possibility that some proteins do not bind on their own, but instead are associated with other target proteins. The purified forms of phosphodiesterase, rGARP2, T_α , T_β and holo- $T_{\alpha\beta}$ tended to undergo hydrophobic

|| Present addresses: Biomedizinische Forschung, Merck AG, 64271 Darmstadt, Germany (R.K.); Institut für Zoologie, Universität Mainz, 55099 Mainz, Germany (U.W., C.B.)

interactions with the column matrix (data not shown). Small amounts of detergents reduced non-specific binding of phosphodiesterase and rGARP2 to control columns, while preserving binding to R3 (Fig. 2c); therefore, phosphodiesterase and rGARP2 interact directly with R3. In contrast, detergent abolished binding of T_{α} and $T_{\alpha\beta\gamma}$ to both R3 and control columns (shown for T_{α} in Fig. 2c). This result, along with the gel-filtration experiment (see below), implies that transducin interacts, at most, only weakly with GARP repeats. The purified cGMP-gated channel, consisting of α - and β -subunits⁴, bound to R3 but not to the control peptide (Fig. 2c). A dodecyl maltoside extract of ROS membranes¹² highly enriched in guanylate cyclase also contained tubulin and actin, but no ABCR or cGMP-gated channel or rhodopsin (ref. 13; and data not shown). The guanylate cyclase was completely removed from the extract (Fig. 2c) by the R3 column, along with tubulin and actin (data not shown) but not by a control. This demonstrates that cytoskeletal elements participate in a complex involving guanylate cyclase¹³ and GARPs. Purified target proteins also bound to R2, but not to 'scrambled' R2 (having the same amino-acid composition as R2, but in random sequence; data not shown).

We used a complementary approach to demonstrate a high-

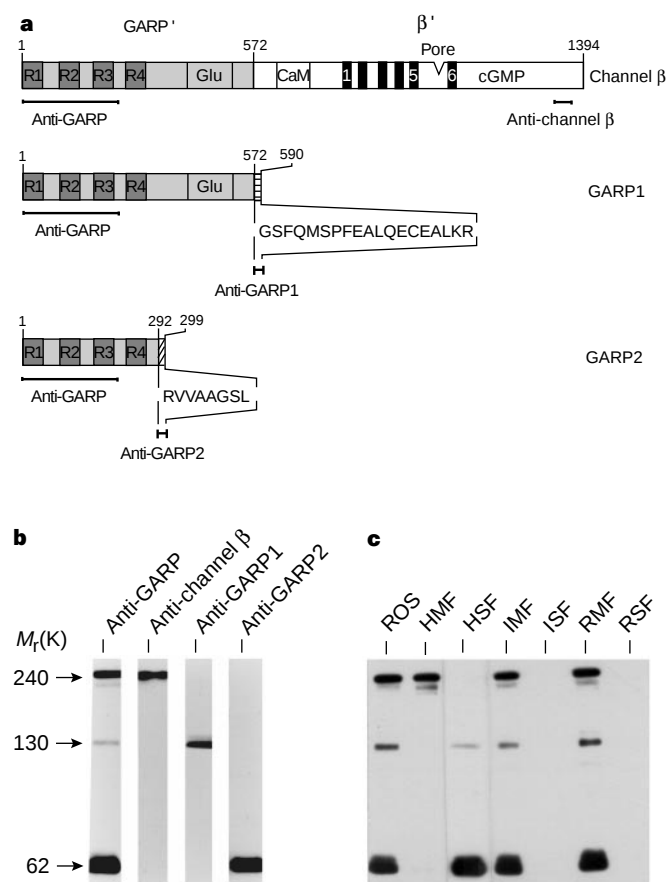
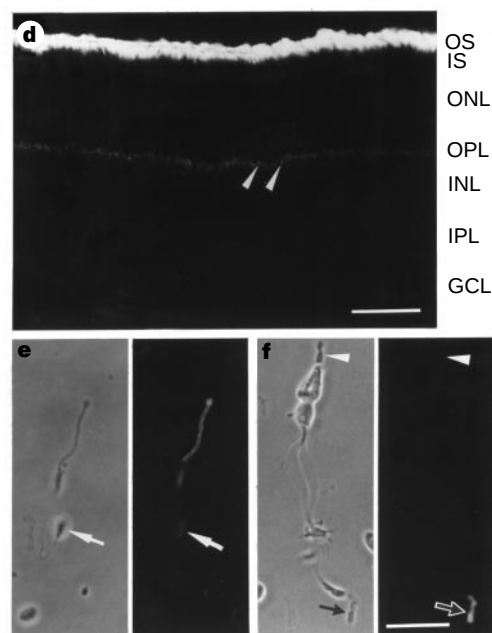


Figure 1 Structural features of GARPs and localization to rod photoreceptors. **a**, Diagram of GARP1, GARP2 and the β -subunit of the cGMP-gated channel, highlighting the four repeats (R1-R4), the glutamic-acid-rich region (Glu), a calmodulin-binding site (CaM), the transmembrane segments (1-6), the pore region and the binding site for cGMP. **b**, Immunoblots showing the size and relative abundance of the three GARP polypeptides in ROS. Western blot (5 μ g rhodopsin) was probed with antibodies specific for GARP1 (anti-GARP1), GARP2 (anti-GARP2), the β subunit (anti-channel β) and all three GARP polypeptides (anti-GARP). The epitopes against which antibodies were directed are shown in **a**. **c**, Membrane association of GARPs. Western blot was probed with the anti-GARP antibody. HMF, hypotonic membrane fraction; HSF, hypotonic soluble fraction; IMF, isotonic membrane fraction; ISF, isotonic soluble fraction; RMF, reconstituted

affinity complex between native GARPs and phosphodiesterase. Soluble ROS proteins were fractionated by gel filtration and the elution profiles of phosphodiesterase and GARPs were analysed by western blotting (Fig. 2d). In a cytosolic extract of illuminated ROS, most GARP2 almost perfectly co-elutes with phosphodiesterase in two column fractions that predict a M_r of ~ 230 – 270 K for the complex, whereas a smaller portion of GARP2 elutes in column fractions compatible with a monomer (~ 30 K) and a dimer (~ 60 K). For an extract prepared in the dark, the elution profile is shifted towards the monomeric/dimeric forms of GARP2 (Fig. 2d). These results show that GARP2 is in binding equilibrium with a high-affinity phosphodiesterase/GARP2 complex and that active phosphodiesterase displays a higher affinity for GARP2 than does holo-phosphodiesterase.

Therefore, we examined whether GARP2 affects phosphodiesterase activity. Recombinant GARP2 up to concentrations of 5 μ M did not affect the basal activity of purified phosphodiesterase (Fig. 3a). Phosphodiesterase activation by purified T_{α} -GTP γ S (molar ratio phosphodiesterase: T_{α} = 1 : 8, hydrolysis turnover 200–320 cGMP molecules s^{-1}) was inhibited by rGARP2 to almost basal levels (10–20%), with 50% inhibitory concentration of roughly 10 nM (Fig. 3a). In contrast, the activity of trypsinized phos-



g

R1	1-	MLGWV	---	Q	R	V	L	P	Q	P	P	-13		
R2	100-	VLTWLRK	G	V	E	K	V	V	P	Q	P	A	-116	
R3	167-	LLRWFE	Q	N	L	E	K	M	L	P	Q	P	P	-183
R4	255-	LMAWIL	H	R	L	E	M	A	L	P	Q	P	V	-271

membrane fraction; RSF, reconstituted soluble fraction. Samples contained roughly 5 μ g rhodopsin. **d**, Vertical section of the bovine retina stained with antibody Anti-GARP. Rod outer segments (OS) and rod spherules in the outer plexiform layer (OPL, arrowheads) are labelled. GCL, ganglion cell layer; INL, inner nuclear layer; IPL, inner plexiform layer; IS, inner segment; ONL, outer nuclear layer. **e**, Isolated rod photoreceptor shown in phase contrast (left) and under epifluorescence (right). **f**, No staining is found in an isolated cone; the arrowhead marks the cone OS, the arrow points to a nearby rod OS that is labelled. Scale bars: 50 μ m (**d**); 20 μ m (**e**, **f**). **g**, Amino-acid alignment of R1-R4 from bovine GARPs. Black boxes indicate amino-acid identities; grey boxes show conservative substitutions.

phosphodiesterase was not inhibited; if anything, activity increased slightly at high rGARP2 concentrations (Fig. 3a), suggesting that inhibition by GARP2 requires the phosphodiesterase γ -subunit that is degraded by trypsin and/or T_{α} -GTP γ S. We obtained similar results with partially purified native GARP2. In rod photoreceptors, light activates phosphodiesterase on the surface of the disc membrane. To mimic this situation, we used a near-infrared light-scattering signal¹⁴ to monitor the activation of membrane-bound

phosphodiesterase by light. The 'phosphodiesterase signal' was progressively attenuated by increasing amounts of rGARP2 (Fig. 3b). The fraction of the signal that could be maximally suppressed varied between protein preparations, but was usually 50–80%. Thus, rGARP2 inhibits both soluble and membrane-bound phosphodiesterase.

The ABCR^{8,9} (also called the rim protein¹⁰) and guanylate cyclase¹⁵ lie at the disc margins. The identification of R1–R4 as

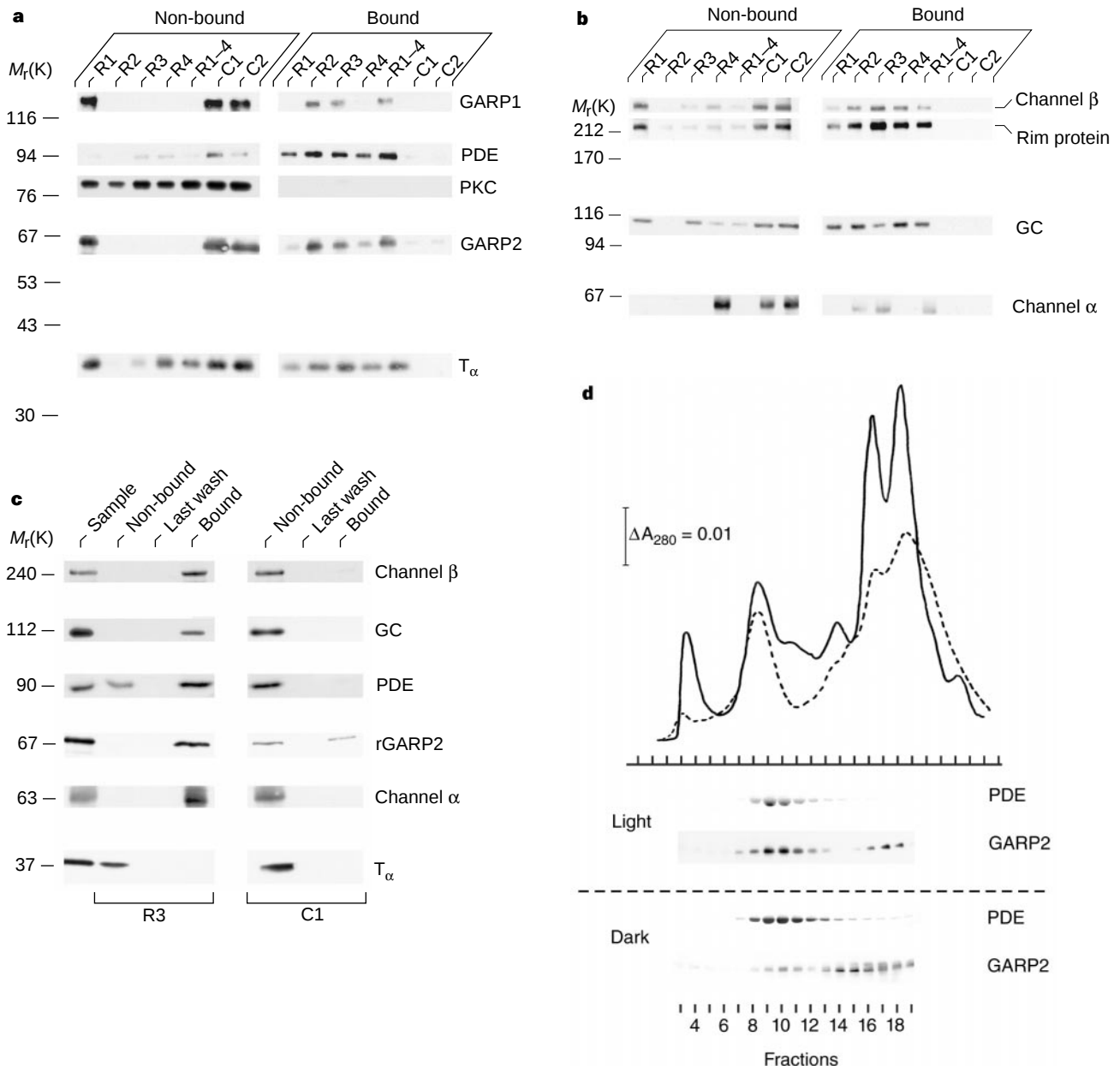


Figure 2 Association of rod outer segment proteins with GARPs. Analysis of binding to GARP repeats R1–R4 by affinity chromatography with soluble proteins (**a**), membrane proteins (**b**) and purified components (**c**). **a**, Western-blot analysis of bound and non-bound column fractions of cytosolic proteins using the following antibodies: polyclonal anti-GARP; polyclonal α' -III, specific for phosphodiesterase (a gift from J. Beavo) monoclonal 3, specific for the α -subunit of protein kinase C (PKC; Transduction Laboratories); and G_{α} -11 (K-20), specific for T_{α} (Santa Cruz Biotechnology). Columns either carried a single repeat (R1, R2, R3, R4) or a mixture of the four repeats (R1–R4), or, as controls, an unrelated peptide (C1) or no peptide (empty column) (C2). **b**, Western-blot analysis bound and non-bound column fractions of integral membrane proteins using the following antibodies: polyclonal anti-channel β (PPc32K in ref. 4), specific for the β -subunit;

polyclonal FPc39K, specific for the α -subunit of the channel; polyclonal ABCR2374, specific for the ABC transporter (a gift from J. Nathans); and polyclonal PPc3, specific for guanylate cyclase (GC). **c**, Affinity chromatography of purified proteins with immobilized R3 and control columns (see Methods). Western-blot signals of the sample input that had been applied to the column (Sample), of the non-bound fraction (Non-bound), the last-wash fraction (Last wash) and the fraction eluted by SDS buffer (Bound). **d**, Gel-filtration chromatography on a Superdex-200 column of dark and light extracts of cytosolic proteins. Upper part, protein elution profile of light (continuous) and dark (dashed) extract (0.5-ml fractions) monitored at 280 nm. Lower part, proteins were visualized with Coomassie (PDE) and by western blot using anti-GARP antibody (GARP2).

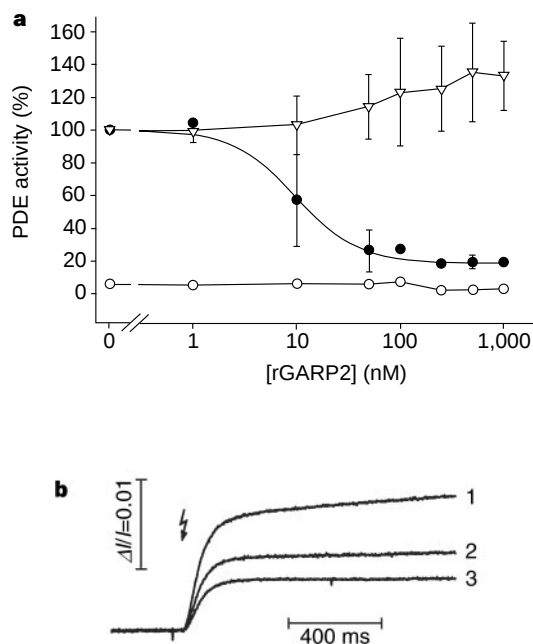


Figure 3 Inhibition of phosphodiesterase (PDE) by rGARP2. **a**, Normalized PDE activity (%) as a function of rGARP2 concentration. Holo-PDE was activated either by 1.5×10^{-7} M T_{α} -GTP γ S (●) or by trypsinization (∇); ○, effect of rGARP2 on basal activity. PDE activity (turnover number s^{-1}) was normalized to the activity without the addition of rGARP2. Basal activity of holo-PDE was normalized to trypsinized PDE activity. Data represent means of at least three determinations $\pm$ s.d. **b**, Near-infrared light-scattering monitor of PDE activation by a light flash and inhibition by rGARP2; trace 1, 3 μ M rhodopsin, 0.35 μ M transducin, 0.2 μ M PDE, 500 μ M GTP, no rGARP2; trace 2, plus 0.2 μ M rGARP2; trace 3, plus 0.6 μ M rGARP2.

protein-recognition domains raises the possibility that GARPs act *in vivo* to assemble a protein complex at the disc rim, in close proximity to the cGMP-gated channel in the plasma membrane¹⁶. If GARP2 interacts with one or several of these membrane proteins, it should be predominantly located at the disc margin or the plasma membrane. We studied the subcellular distribution of GARPs by electron microscopy (Fig. 4). Silver-enhanced immunogold labelling with the anti-GARP antibody revealed that GARPs are localized exclusively at the circumference and incisures of rods in longitudinal and transverse sections of bovine ROS (Fig. 4a–c). Rod inner segments and cone outer segments were not labelled (Fig. 4a). Mammalian rods usually have a single incisure per disc¹⁷ and incisures of adjacent discs are aligned for several micrometres along the outer segment. Several of these longitudinal clefts are laterally displaced¹⁸. Therefore, decoration of incisures similar to the pattern shown in Fig. 4c is typical for mammalian rods and is also observed for the ABCR⁸. These results demonstrate that all GARPs are restricted to the small subcellular space between the plasma membrane (GARP' part) and disc rim (GARP1 and GARP2). In contrast, we observed uniform immunogold labelling of retinal phosphodiesterase throughout the outer segment (Fig. 4d). In control experiments, preincubation of the antibody with rGARP2 (1 mg ml⁻¹) entirely abolished ROS labelling, whereas preincubation with an unrelated recombinant protein (centrin 1) had no effect (data not shown).

In conclusion, we provide evidence that soluble GARPs are confined to the rim region and incisures of discs. Although the electron micrographs cannot prove a physical interaction, it is

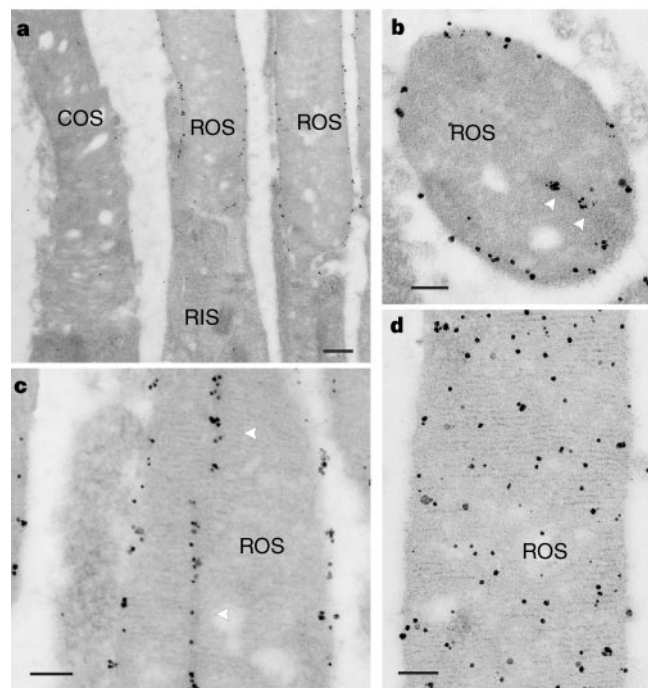
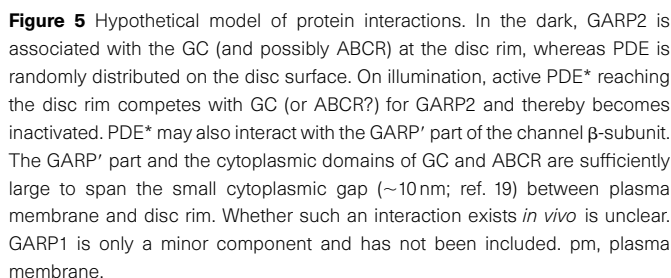


Figure 4 Immunoelectron microscopic localization of GARPs and PDE in photoreceptor cells. **a**, Longitudinal section through bovine photoreceptor cells. Silver-enhanced immunogold labelling with anti-GARP is restricted to peripheral regions of ROS. Rod inner segments (RIS) and cone outer segments (COS) are not labelled. **b**, Transverse section; **c**, longitudinal section through bovine ROS. Silver-enhanced immunogold labelling with anti-GARP is present at peripheral regions and in incisures (arrowheads) of ROS. **d**, Electron micrograph of a longitudinal section through a ROS, illustrating silver-enhanced immunogold labelling of retinal PDE. The entire outer segment is labelled. Scale bars: **a**, 500 nm; **b–d**, 200 nm.

plausible that GARPs are confined to the rim by binding to guanylate cyclase and the ABCR, which also occur at the disc margins^{8,10,15} (Fig. 5). The GARP' part of the channel β -subunit, in principle, may also interact with the large cytoplasmic domains of guanylate cyclase and ABCR across the narrow gap between plasma and disc membrane (~ 10 nm; ref. 19). Co-immunoprecipitation experiments failed to demonstrate the existence of such a protein complex unequivocally (unpublished). However, solubilization of membrane proteins, in particular of guanylate cyclase, required high detergent concentrations that might have disrupted a complex between components of the plasma and disc membrane.

What is the function of GARPs? Active phosphodiesterase is turned off by the endogenous GTPase activity of T_{α} -GTP involving RGS9 (ref. 20)—a member of the family of GAP proteins—and the β -subunit of type-5 G protein²¹. The powerful inhibition of active phosphodiesterase by rGARP2 suggests an independent mechanism of phosphodiesterase inactivation that might be important for termination and adaptation of the light response: active phosphodiesterase molecules that reach the disc rim by diffusion will be deactivated by GARP2 or the GARP' part of the channel. An intriguing scenario is that the two inactivation mechanisms operate in different light regimes. Sequestration by GARPs may become the dominant inactivation mechanism at high light intensities, when it is highly probable that a significant fraction of active phosphodiesterase is reaching the disc margins²². In this respect, GARPs may organize an 'adaptational' signalling complex at the disc rim that downregulates the high cGMP turnover during daylight, when rod function is saturated. □



Western blot and immunohistochemistry. Polyclonal antibodies were produced and purified as described previously⁴. Antibody anti-GARP1 was raised against a peptide from the C terminus of GARP1 (amino acids 572–590); anti-GARP2 against a peptide from the C terminus of GARP2 (290–299); anti-GARP against a glutathione S-transferase fusion protein of the N-terminal sequence (5–221) common to GARP1, GARP2 and the GARP' part of the rod cGMP-gated channel β -subunit. Western blotting was performed as described previously^{4,23}, as were immunohistochemistry and dissociation of interact

Gel filtration. Supernatants of hypotonically washed ROS were fractionated by gel-filtration chromatography on a Superdex-200 column (Pharmacia). Phosphodiesterase and GARPs were detected in elution fractions (0.5 ml) by western blotting using Coomassie staining (phosphodiesterase) and the anti-GARP antibody, respectively.

Affinity-column chromatography. The amide form of peptides was synthesized, purified and characterized as described elsewhere²⁴. The sequences of peptides were: R1, MLGWVQRVLPQPPG; R2, GWVLTWLRKGVKEVVPQPAH; R3, PWLLRWFEQNLEKMLPQPPK; R4, ARLMAWILHRLEMALPQPV; and control peptides (C1) GEGRLKVLQE or 'scrambled' R2 (GHWA-VPLQTPWVYLVRKKEGV). Peptides were immobilized to activated CH

Phosphodiesterase assay. Purified phosphodiesterase was preincubated in a final volume of 50 μ l with T_{α} -GTP γ S and various rGARP2 concentrations in a buffer containing 200 mM NaCl, 10 mM HEPES pH 7.4, 2 mM DTT, 2 mM $MgCl_2$ for 10 min at 25 $^{\circ}C$. The concentration of holo-phosphodiesterase was 1.75×10^{-8} M and that of trypsinized phosphodiesterase was 2×10^{-9} M. The reaction was started by adding 500 μ M cGMP and was stopped after 20 s by adding 50 μ l ice-cold 100 mM EDTA and boiling for 5 min. The cGMP and 5'-GMP concentrations were analysed by HPLC using a nucleotide separation and quantification system as described elsewhere²⁶.

Real-time light-scattering assay. Changes in light-induced near-infrared light scattering of membrane suspensions that have been reconstituted with purified transducin and phosphodiesterase were measured as described previously¹⁴.

Electron microscopy. Immunoelectron microscopy was performed with LR White sections of isolated bovine retinae as described^{27,28} previously. Nanogold (Nanoprobes, Stony Brook) labelling of ultra-thin sections was silver-enhanced as described²⁹ and examined in a Zeiss EM 910 electron microscope.

1. Tsunoda, S. *et al.* A multivalent PDZ-domain protein assemblies signalling complexes in a G-protein-coupled cascade. *Nature* **388**, 243–249 (1997).
2. Zuker, C. S. & Ranganathan, R. The path to specificity. *Science* **283**, 650–651 (1999).
3. Sugimoto, Y., Yatsunami, K., Tsujimoto, M., Khorana, H. G. & Ichikawa, A. The amino acid sequence of a glutamic acid-rich protein from bovine retina as deduced from the cDNA sequence. *Proc. Natl Acad. Sci. USA* **88**, 3116–3119 (1991).
4. Körschen, H. G. *et al.* A 240 kDa protein represents the complete β subunit of the cyclic nucleotide-gated channel from rod photoreceptor. *Neuron* **15**, 627–636 (1995).
5. Ardell, M. D. *et al.* cDNA, gene structure, and chromosomal localization of human GAR1 (CNCG3L), a homolog of the third subunit of bovine photoreceptor cGMP-gated channel. *Genomics* **28**, 32–38 (1995).
6. Colville, C. A. & Molday, R. S. Primary structure and expression of the human β -subunit and related proteins of the rod photoreceptor cGMP-gated channel. *J. Biol. Chem.* **271**, 32968–32974 (1996).
7. Grunwald, M. E., Yu, W.-P., Yu, H.-H. & Yau, K.-W. Identification of a domain on the β -subunit of the rod cGMP-gated cation channel that mediates inhibition by calcium–calmodulin. *J. Biol. Chem.* **273**, 9148–9157 (1998).
8. Illing, M., Molday, L. L. & Molday, R. S. The 220-kDa rim protein of retinal rod outer segment is a member of the ABC transporter superfamily. *J. Biol. Chem.* **272**, 10303–10310 (1997).
9. Allikmets, R. *et al.* A photoreceptor cell-specific ATP-binding transporter gene (*ABCR*) is mutated in recessive Stargardt macular dystrophy. *Nature Genet.* **15**, 236–246 (1997).
10. Papermaster, D. S., Schneider, B. G., Zorn, M. A. & Kraehenbuhl, J. P. Immunocytochemical localization of a large intrinsic membrane protein to the incisures and margins of frog rod outer segment disks. *J. Cell Biol.* **78**, 415–425 (1978).
11. Montell, C. TRP trapped in fly signaling web. *Curr. Opin. Neurobiol.* **8**, 389–397 (1998).
12. Koch, K.-W. & Lambrecht, H.-G. in *Signal Transduction in Photoreceptor Cells* (eds Hargrave, P. A., Hofmann, K. P. & Kaupp, U. B.) 259–267 (Springer, Berlin, 1992).
13. Schrem, A., Lange, C., Beyermann, M. & Koch, K.-W. Identification of a domain in guanylyl cyclase-activating protein 1 that interacts with a complex of guanylyl cyclase and tubulin in photoreceptors. *J. Biol. Chem.* **274**, 6244–6249 (1999).
14. Heck, M. & Hofmann, K. P. G-protein-effector coupling: A real-time light-scattering assay for transducin-phosphodiesterase interaction. *Biochemistry* **32**, 8220–8227 (1993).
15. Liu, X. *et al.* Ultrastructural localization of retinal guanylate cyclase in human and monkey retinas. *Exp. Eye Res.* **59**, 761–768 (1994).
16. Cook, N. J., Molday, L. L., Reid, D., Kaupp, U. B. & Molday, R. S. The cGMP-gated channel of bovine rod photoreceptors is localized exclusively in the plasma membrane. *J. Biol. Chem.* **264**, 6996–6999 (1989).
17. Pedler, C. M. & Tilly, R. The fine structure of photoreceptor discs. *Vision Res.* **7**, 829–836 (1967).
18. Eckmiller, M. S. & Toman, A. Association of kinesin with microtubules in diverse cytoskeletal systems in the outer segments of rods and cones. *Acta Anat.* **162**, 133–141 (1998).
19. Roof, D. J. & Heuser, J. E. Surfaces of rod photoreceptor disk membranes: Integral membrane components. *J. Cell Biol.* **95**, 487–500 (1982).

20. He, W., Cowan, C. W. & Wensel, T. G. RGS9, a GTPase accelerator for phototransduction. *Neuron* **20**, 95–102 (1998).
21. Makino, E. R., Handy, J. W., Li, T. & Arshavsky, V. Y. The GTPase activating factor for transducin in rod photoreceptors is the complex between RGS9 and type 5 G protein β subunit. *Proc. Natl Acad. Sci. USA* **96**, 1947–1952 (1999).
22. Pugh, E. N. Jr & Lamb, T. D. Amplification and kinetics of the activation steps in phototransduction. *Biochim. Biophys. Acta* **1141**, 111–149 (1993).
23. Bönigk, W. et al. Rod and cone photoreceptor cells express distinct genes for cGMP-gated channels. *Neuron* **10**, 865–877 (1993).
24. Zoche, M., Bienert, M., Beyermann, M. & Koch, K.-W. Distinct molecular recognition of calmodulin-binding sites in the neuronal and macrophage nitric oxide synthases: A surface plasmon resonance study. *Biochemistry* **35**, 8742–8747 (1996).
25. Hsu, Y.-T. & Molday, R. S. Modulation of the cGMP-gated channel of rod photoreceptor cells by calmodulin. *Nature* **361**, 76–79 (1993).
26. Lambrecht, H.-G. & Koch, K.-W. A 26 kd calcium binding protein from bovine rod outer segments as modulator of photoreceptor guanylate cyclase. *EMBO J.* **10**, 793–798 (1991).
27. Wolfrum, U. Centrin in the photoreceptor cells of mammalian retinae. *Cell Motil. Cytoskel.* **32**, 55–64 (1995).
28. Wolfrum, U., Liu, X., Schmitt, A., Udovichenko, I. P. & Williams, D. S. Myosin VIIa as a common component of cilia and microvilli. *Cell Motil. Cytoskel.* **40**, 261–271 (1998).
29. Danscher, G. Localization of gold in biological tissue. A photochemical method for light and electron microscopy. *Histochemistry* **71**, 81–88 (1981).

Acknowledgements. We thank J. Beavo, J. Nathans, I. Weyand and D. Weitz for antibodies, A. Eckert for preparing the manuscript, I. Weyand and R. Seifert for critical reading of the manuscript, and E. Pugh and R. Cote for discussions. This work was supported by grants from the EU (U.B.K.), the Deutsche Forschungsgemeinschaft (M.B., K.P.H., K.-W.K., U.B.K., U.W.) and the FAUN-Stiftung (U.W.).

Correspondence and requests for materials should be addressed to U.B.K. (e-mail: a.eckert@fz-juelich.de).

Increased affiliative response to vasopressin in mice expressing the V_{1a} receptor from a monogamous vole

Larry J. Young*, Roger Nilsen*, Katrina G. Waymire†, Grant R. MacGregor† & Thomas R. Insel*

* Department of Psychiatry and Behavioural Sciences and † Center for Molecular Medicine, Emory University, Atlanta, Georgia 30322, USA

Arginine vasopressin influences male reproductive and social behaviours in several vertebrate taxa¹ through its actions at the V_{1a} receptor in the brain. The neuroanatomical distribution of vasopressin V_{1a} receptors varies greatly between species with different forms of social organization^{2,3}. Here we show that centrally administered arginine vasopressin increases affiliative behaviour in the highly social, monogamous prairie vole, but not in the relatively asocial, promiscuous montane vole. Molecular analyses indicate that gene duplication and/or changes in promoter structure of the prairie vole receptor gene may contribute to the species differences in vasopressin-receptor expression. We further show that mice that are transgenic for the prairie vole receptor gene have a neuroanatomical pattern of receptor binding that is similar to that of the prairie vole, and exhibit increased affiliative behaviour after injection with arginine vasopressin. These data indicate that the pattern of V_{1a} -receptor gene expression in the brain may be functionally associated with species-typical social behaviours in male vertebrates.

Arginine vasopressin (AVP) and its non-mammalian homologue, arginine vasotocin (AVT), are involved in many species-typical, male social behaviours, including communication^{4–7}, aggression^{4–6}, sexual behaviour⁷ and, in monogamous species, pairbonding and paternal care⁸. Pharmacological studies using selective antagonists indicate that the behavioural effects of AVP are mediated by the V_{1a} receptor^{8,9}. The V_{1a} receptor is a member of a family of evolutionarily related receptors for the neurohypophysial peptides, arginine vasopressin and oxytocin, which includes the V_{1a} , V_{1b} , V_2 and oxytocin receptors, all of which are G-protein-coupled proteins comprising seven transmembrane domains¹⁰. The V_{1a}

receptor is exceptional in that its pattern of expression in the brain is phylogenetically plastic, with the neuroanatomical distribution of this receptor being unique in virtually every species examined¹⁰.

Voles provide a useful model to test the relationship between receptor expression patterns and social behaviour. Prairie voles (*Microtus ochrogaster*) are highly affiliative, biparental and monogamous; whereas montane voles (*Microtus montanus*) are relatively asocial, non-paternal, and promiscuous¹¹. In the male prairie vole, cohabitation and/or mating with a female stimulates both the central release of AVP^{12,13} and the development of a pairbond and paternal care^{8,12}. Central blockade of the vasopressin receptors during mating with a selective V_{1a} antagonist prevents the normal development of pairbonds between mates and paternal care in male prairie voles^{8,14}. Prairie and montane voles have strikingly different distributions of V_{1a} -receptor binding in the brain (Fig. 1a). We tested the hypothesis that these differences in V_{1a} -receptor binding patterns would confer species differences in behavioural response to AVP. Using an ovariectomized stimulus female placed in one side of a two-chambered arena, 2 ng of AVP delivered by intracerebroventricular (i.c.v.) injection to male prairie voles resulted in a significant increase in affiliative behaviour (olfactory investigation and grooming) toward the stimulus female ($P < 0.01$) (Fig. 1b). In contrast, the identical treatment had no effect on affiliative behaviour, compared with control injections, in male montane voles.

To investigate the mechanisms of the species differences in regional gene expression, we compared the structures of the prairie and montane vole V_{1a} -receptor genes, including the 5' flanking regions (Fig. 2a). The V_{1a} -receptor gene encodes a protein of 420 amino-acid residues that is 99% homologous between the vole species. The binding kinetics of the receptor protein and its second messenger coupling are identical in both species³. Both 5' and 3' rapid amplification of complementary DNA ends (RACE) were used to identify the transcriptional boundaries of the gene (data not shown). The prairie vole V_{1a} messenger RNA is composed of a 232-base-pair (bp) 5' untranslated region (UTR), a 1,260-bp coding region followed by a 36-bp 3' UTR. No CAAT or TATA enhancer

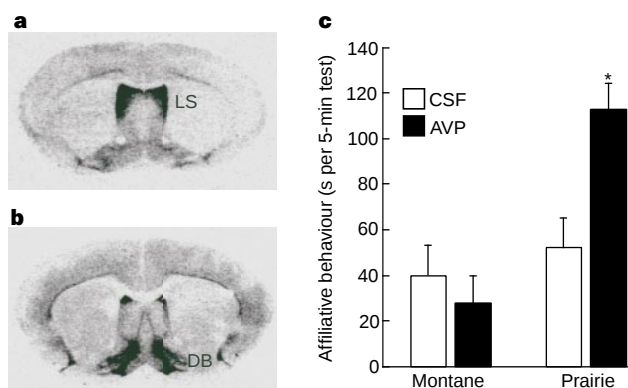


Figure 1 Montane and prairie voles differ both in V_{1a} -receptor binding pattern and behavioural response to arginine vasopressin. Receptor autoradiography⁶ illustrates the different patterns of V_{1a} -receptor binding in the brains of **a**, the non-monogamous montane vole, and **b**, the monogamous prairie vole. Note the high intensity of binding in the lateral septum (LS) of the montane vole but not the prairie vole, and in the diagonal band (DB) of the prairie vole but not the montane vole. Similar differences exist throughout the brain. **c**, Male prairie but not montane voles exhibit elevated levels of affiliative behaviour after vasopressin is administered directly into the brain (two-way ANOVA: species effect, $F(1,27) = 10.3$, $P < 0.01$; * Fisher's LSD post-hoc test, $P < 0.01$ compared with CSF-treated prairie voles).

elements were found near the transcriptional initiation site. Northern analysis using poly (A)⁺ RNA from brain demonstrate that the transcript length is similar in prairie and montane voles (data not shown). Both species have a 2.5-kilobase (kb) intron located between the sixth and seventh transmembrane domains. In contrast to the similarities in the coding sequences, the 5' flanking region of the *V_{1a}* gene displays marked differences between the species. In the gene for the prairie vole *V_{1a}*-receptor, this region contains a 428-bp sequence that is rich in repetitive di- and tetra-nucleotide sequences, or microsatellite DNA, between -1,151 and -723 relative to the transcription start site. This sequence is not found in the montane vole *V_{1a}* gene and sequences on either side of the expansion are contiguous in the montane vole gene.

Using polymerase chain reaction (PCR) primers positioned on either side of the expansion, we found that the *V_{1a}*-receptor gene of the monogamous pine vole (*Microtus pinetorum*), which has a prairie vole-like pattern of *V_{1a}*-receptor expression³, has a similar 5' flanking sequence to that of the prairie vole gene (Fig. 2b). In contrast, the non-monogamous meadow vole (*Microtus pennsylvanicus*), which has a montane-vole-like pattern of receptors in the brain, did not have this sequence. In addition to the change in the 5' flanking region, two distinct *V_{1a}* loci were isolated from prairie vole genomic libraries, indicating that the *V_{1a}*-receptor gene from the prairie vole has been duplicated (Fig. 2a). Southern blot analysis of genomic DNA from both prairie and montane voles confirms the duplication (not shown). The second *V_{1a}*-receptor locus has a 589-bp 5' truncated fragment of an L1 long interspersed nuclear element (LINE) just upstream of the expansion. The LINE, a mobile genetic element or retrotransposon, inserted in the centre of a 10-bp palindrome (TTCTT/AAGAA). A deletion of a single base in the sequence encoding the third intracellular loop of the receptor results in a frame shift and premature stop.

Microsatellite DNA sequences in close proximity to transcription initiation sites can alter transcription^{15,16}, presumably by altering the structural environment of the region containing the *cis* regulatory sequences. Repetitive tetranucleotide repeats act as transcriptional enhancers in the human tyrosine hydroxylase gene¹⁷. In addition to the changes in the 5' flanking region, gene duplication (that is, translocation to a new chromosomal environment) may contribute to the species differences in gene expression. The present data cannot distinguish between these mechanisms.

To determine whether the pattern of neural expression of *V_{1a}* receptors was sufficient to induce differences in social behaviour, we created transgenic mice using a prairie vole *V_{1a}*-receptor mini-gene

containing 2.2 kb of the 5' flanking region, both exons with the 2.5-kb intron, and 2.4 kb of 3' flanking region. Expression of *V_{1a}* receptors in the central nervous system was detected by receptor autoradiography using the specific *V_{1a}*-receptor ligand, [¹²⁵I]-Lin AVP¹⁸ and *in situ* hybridization. In one of the four transgenic lines created, the neuroanatomical pattern of *V_{1a}*-receptor binding was remarkably similar to that of the prairie vole and was consistent for at least four generations (Fig. 3a, b, e). There were no differences in oxytocin-receptor binding between transgenic and non-transgenic mice (Fig. 3c, d). The pattern of *V_{1a}*-receptor mRNA in transgenic mouse brain was similar to that of ligand binding, demonstrating that the change in receptor binding was due to altered transgene expression in the brain (data not shown). The intensity of binding in the transgenic mouse was not elevated in brain regions that normally express *V_{1a}* receptor in the mouse, such as the diagonal band, lateral septum or the lateral hypothalamus (Fig. 3f). The level of expression of the prairie vole *V_{1a}* transgene in the mice was within the range of that found in prairie voles. Thus, any changes in response to AVP should be attributable to the expression of the receptor in new areas of the mouse brain and not simply to increased numbers of receptors in regions that are normally responsive to AVP in the mouse.

In a behavioural model similar to that used in the vole experiments, 2 ng of AVP increased the expression of affiliative behaviour, as measured by olfactory investigation and grooming, in transgenic but not non-transgenic mice ($P < 0.01$) (Fig. 4a), a response similar

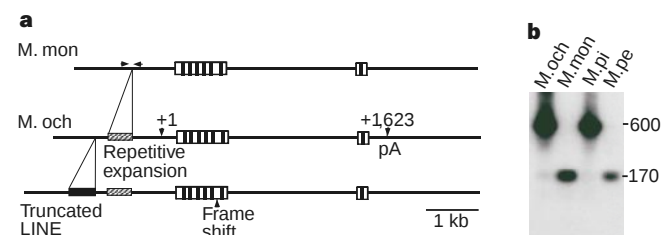


Figure 2 Structure of the *V_{1a}*-receptor gene in voles. *V_{1a}*-receptor genes from montane vole (M. mon) and prairie vole (M. och) were isolated from genomic DNA libraries. Transcription begins at +1 and polyadenylation (pA) occurs at +1,623 of the prairie vole gene. The boxed area within these sites represents the coding region and the vertical bars in the coding region represent the location of the transmembrane domains. Two distinct loci were isolated from prairie vole DNA. The sequences of the genomic clones have been deposited in GenBank (accession number AF069304). **b**, PCR amplification of genomic DNA using primers on either side of the expansion (arrows in **a**) demonstrates the presence of the expansion in monogamous prairie and pine (M. pi) voles (600-bp band), but not in promiscuous montane or meadow (M. pe) voles.

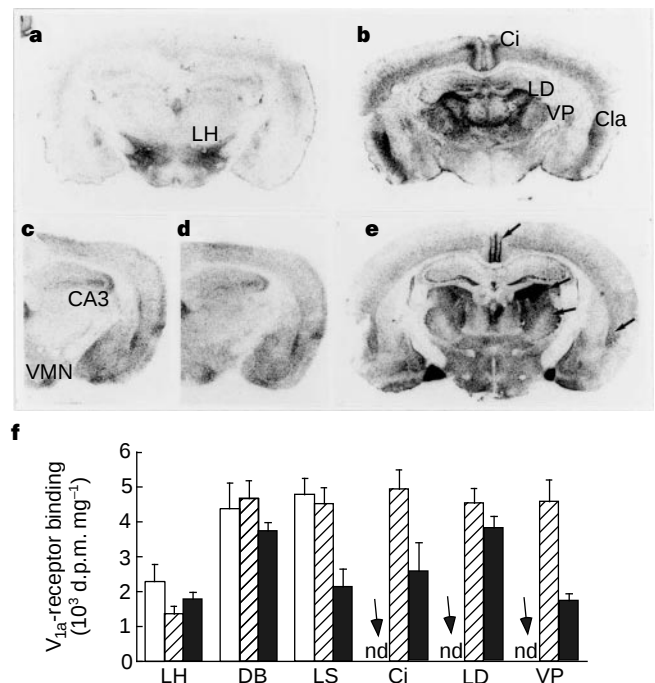


Figure 3 The pattern of *V_{1a}*-receptor binding of mice that are transgenic for the *V_{1a}*-receptor gene from the prairie vole is similar to that of the prairie vole. Receptor autoradiography⁶ illustrates the distribution of *V_{1a}*-receptor binding in wild-type (**a**) and transgenic (**b**) mice. Although *V_{1a}*-receptor binding is altered, oxytocin-receptor binding is identical in wild-type (**c**) and transgenic (**d**) mice. Binding in the cingulate cortex (Ci), laterodorsal (LD) and ventroposterior (VP) thalamic nuclei, claustrum (Cla) and several other regions of transgenic mice is similar to that found in the brain of the prairie vole (**e**). **f**, Quantification of *V_{1a}*-receptor binding in wild-type mice (open bars), transgenic mice (striped bars) and prairie vole (filled bars) ($n = 6-7$ per group) demonstrates that receptor density in transgenic mice is not elevated in areas that normally express receptors in mice (for example, lateral hypothalamus (LH), diagonal band (DB) and lateral septum (LS)); nd indicates no detectable specific binding in wild-type mice.

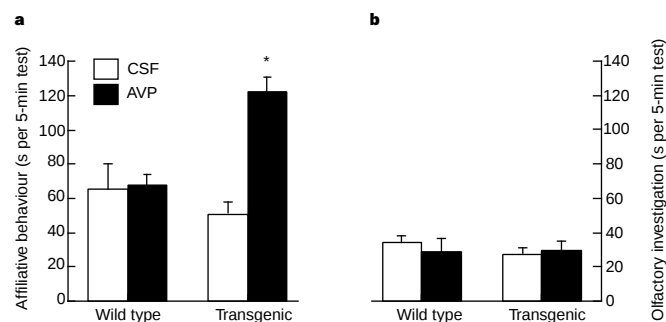


Figure 4 Behavioural response to AVP in transgenic mice. **a**, AVP increases the duration of affiliative behaviours directed towards an ovariectomized female in male transgenic mice with a prairie vole pattern of V_{1a} -receptor expression, but not in male wild-type mice (two-way ANOVA: genotype effect, $F(1,22) = 4.3$, $P < 0.05$; *Fisher's LSD post-hoc test, $P < 0.01$ compared with CSF-treated transgenic animals). **b**, AVP has no effect on olfactory investigation of cotton balls soiled with bedding from an ovariectomized female's cage in either transgenic or non-transgenic males.

to that observed in prairie voles (Fig. 1). This dose of AVP, identical to that used in the vole studies, did not change locomotor or autogrooming behaviour. When not in contact with the female, males of both genotypes actively explored the arena. To determine whether the increase in olfactory investigation was specific for a social stimulus, we tested another set of male mice with the same model but substituted cotton balls that had been scented with lemon extract. No genotype ($F = 0.129$, $P > 0.05$) or treatment ($F = 2.763$, $P > 0.05$) effects were found in olfactory exploration of this new stimulus. In a further test, we used cotton balls scented with bedding from ovariectomized females and still observed no effect of genotype ($F = 0.443$, $P > 0.05$) or treatment ($F = 0.038$, $P > 0.05$) (Fig. 4b). Thus, the increase in olfactory investigation and grooming after AVP injection in the transgenic mice seemed to indicate an increase in affiliative behaviour rather than a non-specific increase in investigation of new olfactory stimuli. At higher doses, transgenic mice, but not non-transgenic littermates, exhibited a dose-dependent increase in autogrooming behaviour after AVP injection (significant at 100 ng, data not shown).

Our results show that centrally administered AVP increases affiliative behaviour in the monogamous male prairie vole, but not in the montane vole. This acute effect of AVP may be essential for the long-lasting effects of AVP in monogamous mammals, such as the formation of pairbonds⁸. The species difference in response to AVP is likely to be due to differences in the distribution of V_{1a} -receptor expression in the brain as (1) vole species with similar social organization have similar receptor patterns³, and (2) transgenic induction of the prairie vole pattern of V_{1a} receptors in mice results in a pro-social response to AVP. Although many genes are likely to be involved in the evolution of complex social behaviours such as monogamy, our data indicate that changes in the expression of a single gene can have an impact on the expression of components of these behaviours, such as affiliation. These observations, together with the species diversity in the neuroanatomical distribution of V_{1a} receptors in the brain, indicate that changes in the pattern of V_{1a} -receptor gene expression in the brain may be a common mechanism to alter the behavioural response to AVP, enabling adaptation to changing socioecological situations. □

Methods

Animals and behaviour tests. Prairie and montane voles used in this study were laboratory-reared animals derived from field-caught specimens. Transgenic mice were generated by pronuclear injection in FVB/N embryos.

The testing arena comprised two standard mouse cages connected by means of a tube to allow free movement between the cages. Ovariectomized, conspecific females were used in all tests to eliminate sexual attractiveness as a possible confounding variable. In both vole and mouse experiments, male subjects were lightly anaesthetized with Metofane and injected i.c.v. with 4 μ l of either artificial cerebrospinal fluid (CSF) or 2 ng arginine vasopressin dissolved in the same volume of CSF. After a 5-min recovery period, the subjects were placed in the empty cage of the testing arena. Behaviour was recorded for 5 min beginning at the moment when the subject first entered the cage containing the stimulus female. This time line was chosen based on preliminary observations of the duration of other AVP effects, such as autogrooming and locomotion. Affiliative behaviour was defined as olfactory investigation or grooming of the stimulus animal. In the experiments with voles, it was necessary to anaesthetize the female during the test as females typically show some aggressive behaviour during the first few minutes of contact with a male. In the experiments with mice, the females were not anaesthetized but were tethered to one side of the cage to restrict movement. In some experiments, female mice were replaced with cotton balls scented with either 10 μ l of lemon extract or with soiled bedding from ovariectomized female cages. All experiments included 5–8 animals per treatment. Duration of affiliative behaviour was analysed using a two-way analysis of variance (ANOVA) with species or genotype and treatment as factors, followed by Fisher's least significant difference (LSD) post-hoc test.

Molecular analysis. Genomic clones were isolated from EMBL3 libraries derived from montane and prairie vole genomic DNA (Clontech), and sequenced on both strands. Amplification of the 5' flanking region of the V_{1a} receptor was done using PCR from approximately 100 ng of genomic DNA isolated from the liver of prairie, montane, meadow and pine voles. Ten per cent of the reaction was electrophoresed on a 1.2% agarose gel, transferred onto nylon membrane, and probed with a ³²P-labelled fragment of the prairie vole V_{1a} -receptor gene containing the amplified region.

Received 7 May; accepted 20 May 1999.

- Young, L. J., Wang, Z. & Insel, T. R. Neuroendocrine bases of monogamy. *Trends Neurosci.* **21**, 71–75 (1998).
- Bester-Meredith, J. K., Young, L. J. & Marler, C. A. Species differences in paternal behavior and aggression in *Peromyscus* and their associations with vasopressin immunoreactivity and receptors. *Horm. Behav.* (in the press).
- Insel, T. R., Wang, Z. & Ferris, C. F. Patterns of brain vasopressin receptor distribution associated with social organization in microtine rodents. *J. Neurosci.* **14**, 5381–5392 (1994).
- Ritters, L. V. & Panksepp, J. Effects of vasotocin on aggressive behavior in male Japanese quail. *Ann. NY Acad. Sci.* **807**, 478–480 (1997).
- Ferris, C. F. et al. Vasopressin/serotonin interactions in the anterior hypothalamus control aggressive behavior in gold hamsters. *J. Neurosci.* **17**, 4331–4340 (1997).
- Young, L. J., Winslow, J. T., Nilsen, R. & Insel, T. R. Species differences in V_{1a} receptor gene expression in monogamous and non-monogamous voles: behavioral consequences. *Behav. Neurosci.* **111**, 599–605 (1997).
- Moore, F. L. & Miller, L. J. Arginine vasotocin induces sexual behavior of newts by acting on cells in the brain. *Peptides* **4**, 97–102 (1983).
- Winslow, J., Hastings, N., Carter, C. S., Harbaugh, C. & Insel, T. A role for central vasopressin in pair bonding in monogamous prairie voles. *Nature* **365**, 545–548 (1993).
- Ferris, C. F. & Potegal, M. Vasopressin receptor blockade in the anterior hypothalamus suppresses aggression in hamsters. *Physiol. Behav.* **44**, 235–239 (1988).
- Barberis, C. & Tribollet, E. Vasopressin and oxytocin receptors in the central nervous system. *Crit. Rev. Neurobiol.* **10**, 119–154 (1996).
- Shapiro, L. E. & Dewsbury, D. A. Differences in affiliative behavior, pair bonding, and vaginal cytology in two species of vole (*Microtus ochrogaster* and *M. montanus*). *J. Comp. Psychol.* **104**, 268–274 (1990).
- Bamshad, M., Novak, M. & DeVries, G. J. Cohabitation alters vasopressin innervation and paternal behavior in prairie voles. *Physiol. Behav.* **56**, 751–758 (1994).
- Wang, Z., Smith, W., Major, D. E. & DeVries, G. J. Sex and species differences in the effects of cohabitation on vasopressin messenger RNA expression in the bed nucleus and stria terminalis in prairie voles (*Microtus ochrogaster*) and meadow voles (*Microtus pennsylvanicus*). *Brain Res.* **650**, 212–218 (1994).
- Wang, Z., Ferris, C. F. & DeVries, G. J. Role of septal vasopressin innervation in paternal behavior in prairie voles (*Microtus ochrogaster*). *Proc. Natl Acad. Sci. USA* **91**, 400–404 (1994).
- Choong, C. S., Kempainen, J. A., Zhou, Z. & Wilson, E. M. Reduced androgen receptor gene expression with first exon CAG repeat expansion. *Mol. Endocrinol.* **10**, 1527–1535 (1997).
- Klesert, T. R., Otten, A. D., Bird, T. D. & Tapscott, S. J. Trinucleotide repeat expansion at the myotonin dystrophy locus reduces expression of *DMAHP*. *Nature Genet.* **16**, 402–406 (1997).
- Meloni, R., Albanese, V., Ravassard, P., Treilhou, F. & Mallet, J. A tetranucleotide polymorphic microsatellite, located in the first intron of the tyrosine hydroxylase gene, acts as a transcription regulatory element *in vitro*. *Hum. Mol. Gen.* **7**, 423–428 (1998).
- Barberis, C. et al. Characterization of a novel, linear radiolabeled vasopressin antagonist: an excellent radioligand for vasopressin V_{1a} receptors. *Neuroendocrinol.* **62**, 135–146 (1995).

Acknowledgements. This work was supported by grants from the NIMH and NCRR to L.J.Y. and T.R.I.

Correspondence and requests for materials should be addressed to L.J.Y. (e-mail: lyoun03@emory.edu). The sequences of the genomic clones have been deposited in GenBank (AF069304).

Heparin is essential for the storage of specific granule proteases in mast cells

Donald E. Humphries^{*}, Guang W. Wong[†], Daniel S. Friend[‡], Michael F. Gurish[†], Wen-Tao Qiu[†], Chifu Huang[†], Arlene H. Sharpe[‡] & Richard L. Stevens[†]

^{*} Department of Medicine, Boston University School of Medicine, and VA Medical Center, Boston, Massachusetts 02130, USA

Departments of [†] Medicine and [‡] Pathology, Brigham and Women's Hospital, and Harvard Medical School, Boston, Massachusetts 02115, USA

All mammals produce heparin, a negatively charged glycosaminoglycan that is a major constituent of the secretory granules of mast cells which are found in the peritoneal cavity and most connective tissues. Although heparin is one of the most studied molecules in the body, its physiological function has yet to be determined. Here we describe transgenic mice, generated by disrupting the *N*-deacetylase/*N*-sulphotransferase-2 gene^{1,2}, that cannot express fully sulphated heparin. The mast cells in the skeletal muscle that normally contain heparin lacked metachromatic granules and failed to store appreciable amounts of mouse mast-cell protease (mMCP)-4, mMCP-5 and carboxypeptidase A (mMC-CPA), even though they contained substantial amounts of mMCP-7. We developed mast cells from the bone marrow of the transgenic mice. Although these cultured cells contained high levels of various protease transcripts and had substantial amounts of mMCP-6 protein in their granules, they also failed to express mMCP-5 and mMC-CPA. Our data show that heparin controls, through a post-translational mechanism, the levels of specific cassettes of positively charged proteases inside mast cells.

Heparin is generated by a series of enzymatic modifications of a non-sulphated precursor³. Thus, we speculated that it might be possible to deduce the physiological function of heparin by creating mice that cannot express this mast-cell specific glycosaminoglycan

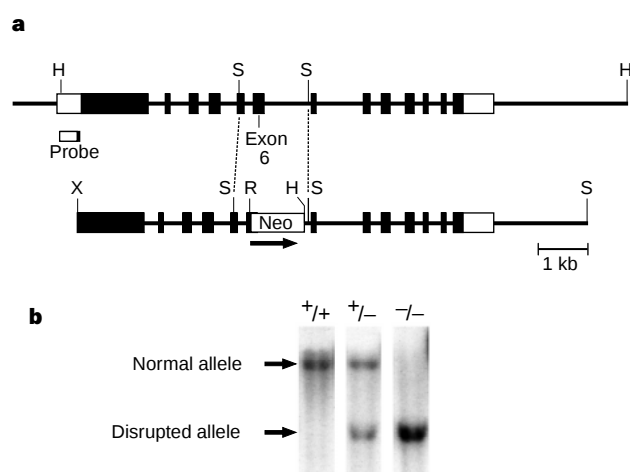


Figure 1 Generation of NDST-2-null mice. **a**, A clone that encodes the entire 13-exon coding region of the NDST-2 gene was isolated from a mouse genomic library to create the indicated targeting vector. H, S, X and R refer to *HindIII*-, *SacI*-, *XhoI*-, and *EcoRI*-susceptible sites, respectively. The 5' and 3' untranslated portions of the gene are indicated in open boxes; the translated portions are indicated in solid boxes. **b**, Blot analysis of *EcoRI*-digested genomic DNA was used to genotype +/+, +/- and -/- mice. The probe used in the genotyping is shown in **a**. Because it resides outside the targeting construct, it can be used to test whether homologous recombination took place.

(GAG) by preventing expression of one of the sulphotransferases used early in its biosynthesis. Two mouse and three human *N*-deacetylase/*N*-sulphotransferases (NDSTs) have been cloned^{1,2,4,5}. Because mouse bone-marrow-derived mast cells (mBMMCs) contain a very high level of the NDST-2 transcript relative to the NDST-1 transcript (data not shown), we concluded that mast cells probably require NDST-2 to synthesize their serglycin-bound heparin chains. We therefore used homologous recombination to disrupt the NDST-2 gene in mice. To this end, we isolated the relevant gene from a 129/Sv mouse genomic library and created the targeting construct shown in Fig. 1a. After transfection and positive selection, an embryonic stem-cell clone was obtained that possessed a disrupted NDST-2 allele. This clone was injected into mouse blastocysts to generate chimaeric mice and then NDST-2-null mice (Fig. 1b).

As assessed by chloroacetate-esterase cytochemistry, mast-cell committed progenitors could home to all analysed tissues of the NDST-2-null mouse (Figs 2, 3). However, the fact that the mast cells in most connective tissues were poorly stained by toluidine blue (Fig. 2b) indicates that disruption of the NDST-2 gene markedly affects heparin expression *in vivo*. The amounts of histamine in the exudates of <4-month-old NDST-2-null mice were $31 \pm 6\%$ (mean $\pm$ range, $n = 2$) of those of wild-type mice. Although the mast cells in the peritoneal cavities of similarly aged NDST-2-null mice contained only small amounts of chloroacetate-esterase activity, the number of chloroacetate-esterase-positive cells at this site was $49 \pm 11\%$ of that in wild-type mice. These data indicate that the amount of histamine in a recognizable peritoneal mast cell in a young wild-type mouse is only $\sim 50\%$ higher than that in a

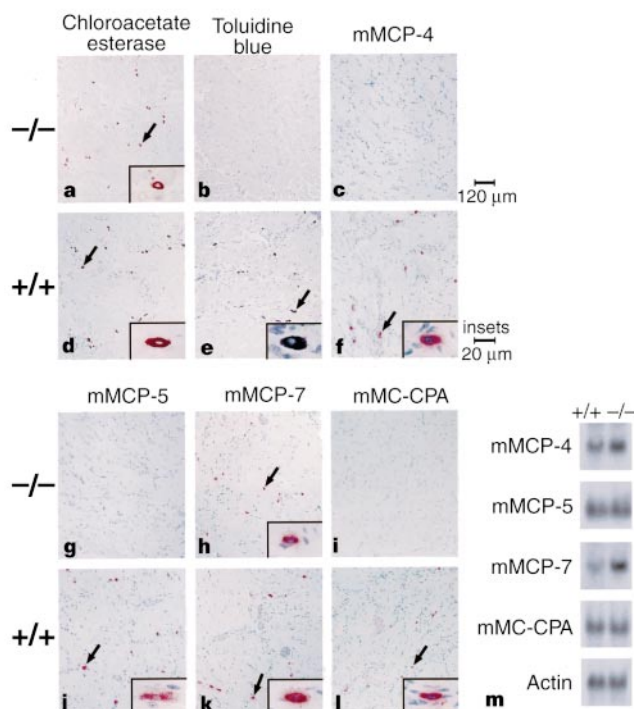


Figure 2 Analysis of mast cells in skeletal muscle and ear. **a-i**, Sections of tongue from NDST-2-null (**a-c**, **g-i**) and wild-type (**d-f**, **j-l**) mice were stained with toluidine blue (**b**, **e**) or with antibodies specific for mMCP-4 (**c**, **f**), mMCP-5 (**g**, **j**), mMCP-7 (**h**, **k**) and mMC-CPA (**i**, **l**). Alternately, tissue sections were evaluated cytochemically for the presence of chloroacetate-esterase-positive cells (**a**, **d**). Arrows indicate mast cells. Insets: higher magnifications of reactive mast cells (**a**, **d-f**, **h**, **j-l**). **m**, Blots containing total RNA from the ear of NDST-2-null mice (right lanes) and their wild-type littermates (left lanes) were analysed with mMCP-4, mMCP-5, mMCP-7, mMC-CPA and β -actin gene-specific probes.

NDST-2-null mouse. However, owing to the difficulty in recognizing poorly granulated mast cells in NDST-2-null mice and to age-dependent variations in the levels of chondroitin sulphate in a peritoneal mast cell, the possibility cannot be ruled out that many of the mast cells found at this site in older mice have substantially less histamine. Essentially no Evans blue dye extravasated into the control ears of wild-type and NDST-2-null mice given antigen alone. In contrast, +++ and ++ passive cutaneous anaphylaxis (PCA) reactions occurred in the ears of wild-type ($n = 4$) and NDST-2-null ($n = 4$) mice, respectively, given both IgE and antigen. It has been thought that the PCA reaction is histamine dependent. Thus, these PCA findings are in agreement with the findings obtained with cultured mast cells (see below) that show that fully sulphated heparin is not essential for histamine expression in every mast cell population.

The mast cells that increase in number in the jejunal mucosal epithelium during helminth infection preferentially express serglycin proteoglycans that contain highly sulphated chondroitins⁶. Therefore, we investigated the consequences of disruption of the NDST-2 gene on this expandable population of mast cells. Not only were mast cells detected in the jejunal mucosa of helminth-infected NDST-2-null mice, but their numbers were ~50% greater than those of their wild-type littermates. As assessed histochemically and cytochemically, the mast cells in the jejunal epithelium of the helminth-infected NDST-2-null mice possessed mature granules (Fig. 3). The one region of the jejunum that lacked metachromatic mast cells was the minor heparin-positive population found in the muscle/submucosa of this tissue in normal mice (Fig. 3, arrows).

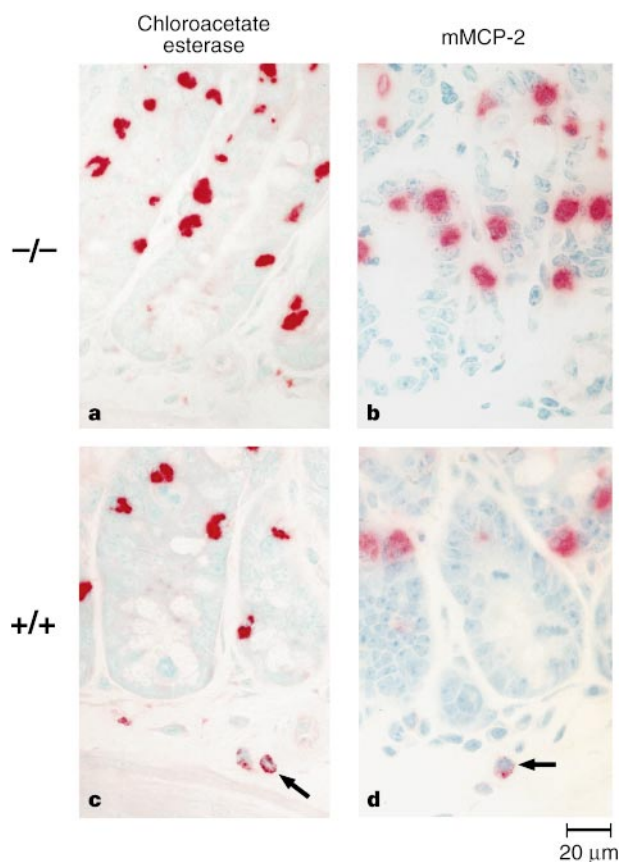


Figure 3 Analysis of mast cells in the jejunum of helminth-infected mice. Sections of jejunum from *T. spiralis*-infected NDST-2-null (a, b) and wild-type (c, d) mice were evaluated cytochemically for the presence of chloroacetate-esterase-positive cells (a, c), or immunohistochemically for the presence of mMCP-2-positive mast cells (b, d). Arrows (c, d) point to muscle/submucosa mast cells that are detectable in wild-type mice but not in NDST-2-null mice.

As well as a variety of negatively charged serglycin proteoglycans, mouse mast cells contain substantial amounts of different combinations of the exopeptidase mMC-CPA and at least 13 distinct serine proteases⁷⁻¹⁴. Whereas the mast cells that increase in number in the jejunal epithelium at the height of *Trichinella spiralis* infection of the BALB/c mouse preferentially express mMCP-2 (refs 8, 9), the constitutive mast cells that are found in normal skin preferentially express mMCP-4, mMCP-5, mMCP-6, mMCP-7 and mMC-CPA^{10-13,15}. Immunohistochemical analysis revealed that the mast cells in the jejunal epithelium of helminth-infected NDST-2-null mice contain large amounts of mMCP-2 (Fig. 3b). The constitutive mast cells in the tongue, skin and ear (Fig. 2m) were also phenotypically mature in terms of their protease gene expression. The mast cells in the tongue contained substantial amounts of the trypsin mMCP-7 (Fig. 2h) and less mMCP-6 (data not shown). However, their levels of mMCP-4, mMCP-5 and mMC-CPA protein were below detection (Fig. 2c, g, i).

Like the *in-vivo*-differentiated mast cells in the peritoneal cavity, the mBMMCs developed from a normal mouse contain mMC-CPA and mMCP-5¹². The two populations of mast cells differ in the nature of the prominent type of GAG attached to serglycin¹⁶. The proteoglycans present in the secretory granules of the mast cells that reside in the peritoneal cavity contain small amounts of chondroitin sulphate and large amounts of heparin. In contrast, the proteoglycans in mBMMCs contain small amounts of heparin and large amounts of chondroitin sulphate. If the protease/proteoglycan interactions in a mast cell are strictly ionic, disruption of heparin biosynthesis should not alter the amounts of mMCP-5 and mMC-CPA in the *in vitro*-differentiated mBMMCs. However, if the protease/proteoglycan interactions are heparin-dependent, there should be substantially less mMCP-5 and mMC-CPA in those mBMMCs developed from NDST-2-null mice. Thus, we generated mBMMCs from NDST-2-null mice to resolve these and other issues. As in the wild type, mast cells could be readily generated *in vitro* from the bone marrow of NDST-2-null mice (Fig. 4). This finding indicates that disruption of the NDST-2 gene does not lead to a decrease in the viability of the mast-cell-committed progenitors in the bone marrow or in their ability to respond to interleukin (IL)-3. The amount of [³⁵S]sulphate transferred from

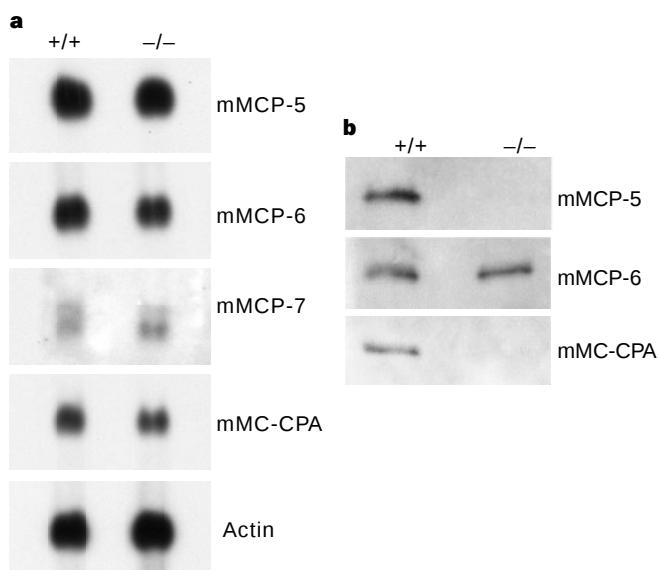


Figure 4 RNA blot and SDS-PAGE/immunoblot analysis of mBMMCs. a, IL-3-dependent mBMMCs developed from NDST-2-null mice (right lanes) and their wild-type littermates (left lanes) were analysed for mMCP-5, mMCP-6, mMCP-7, mMC-CPA and β -actin transcripts. b, Lysates of these cells were also evaluated for mMCP-5, mMCP-6 and mMC-CPA protein.

[³⁵S]PAPS to *N*-desulphated heparin was reduced by at least 75% in the NDST-2-null mBMMCs relative to wild-type mBMMCs. Concomitantly, the granules in the NDST-2-null mBMMCs were poorly stained by toluidine blue (data not shown). Despite the poor granulation, disruption of the NDST-2 gene did not adversely affect histamine expression in these *in vitro*-differentiated mast cells. The histamine content of the null and wild-type mBMMCs was 330 and 320 ng per 10⁶ cells in one experiment and 530 and 440 ng per 10⁶ cells in another, respectively.

Normal mBMMCs markedly increase their biosynthesis of heparin, as well as their granular content of mMCP-5 and mMC-CPA, when co-cultured with fibroblasts¹⁷. We therefore co-cultured mBMMCs developed from NDST-2-null mice with fibroblasts to investigate how disruption of the NDST-2 gene adversely affects heparin biosynthesis and packaging of these two granule proteases. Like normal BALB/c mBMMCs¹⁷, the mBMMCs developed from NDST-2-null mice readily adhered to confluent monolayers of fibroblasts, indicating that they have normal adhesion receptors. The co-cultured mBMMCs developed from the wild-type mice contained safranin-positive granules (Fig. 5a). In contrast, no safranin-positive mast cells could be detected in the co-cultured mBMMCs developed from the NDST-2-null mice. The difference in toluidine blue staining was also marked. Wild-type mBMMCs that had been co-cultured with fibroblasts for 2–4 weeks incorporated 1.4- to 2.1-fold (range, *n* = 3) more [³⁵S]sulphate into proteoglycan than the corresponding mBMMCs developed from NDST-2-null mice. Thus, disruption of the NDST-2 gene caused an overall reduction in sulphated GAGs in these mast cells. We found that 67–74% and 41–51% of the total [³⁵S]proteoglycans produced by the co-cultured NDST-2-null mBMMCs and their wild-type littermates, respectively, were susceptible to chondroitinase ABC. In contrast, 19 to 22% and 46 to 56% of the total [³⁵S]proteoglycans produced by the co-cultured NDST-2-null mBMMCs and their wild-type littermates, respectively, were at least partially susceptible to nitrous acid. DEAE chromatography also revealed that the chondroitinase-ABC-resistant [³⁵S]GAGs produced by the NDST-2-null mBMMCs that had been co-cultured with fibroblasts were less sulphated than those produced by wild-type mBMMCs (data not shown). Moreover, fewer of these heparin-like GAGs were able to bind to an anti-thrombin-III affinity column. The cumulative data indicate that the failure of mast cells to express NDST-2 results in defective formation of structural motifs in the heparin chain, such as the 3-*O*-sulphated glucosamine-containing pentasaccharide that specifically recognizes anti-thrombin III¹⁸.

It has been proposed that heparin is essential for the maturation

of tryptase zymogens¹⁹. We could not determine whether the immunoreactive tryptases (Fig. 2h) in the tongue mast cells of NDST-2-null mice were enzymatically active. Thus, the mBMMCs developed from these animals allowed us to ascertain whether or not heparin was critical for the activation of mMCP-6 and/or mMCP-7 zymogens. The steady-state levels of the transcripts that encode the two tryptases in NDST-2-null mBMMCs were comparable to those in the control mBMMCs developed from their wild-type littermates (Fig. 4a). As assessed by SDS-PAGE/immunoblot analysis and immunohistochemistry, IL-3-developed mBMMCs (Fig. 4b) contained normal levels of mMCP-6 protein. Moreover, the observation that these mast cells contained tryptases in their granules that could cleave a chromogenic substrate indicated that at least some of the expressed tryptases were properly folded and enzymatically active. The data presented in Fig. 4 indicate that mMCP-6 can bind to chondroitin sulphate E. During fibroblast co-culture, mBMMCs switch from synthesis of chondroitin sulphate E to heparin¹⁷. A modest reduction in mMCP-6 content was observed in the co-cultured mBMMCs developed from NDST-2-null mice (Fig. 5b). This probably occurred because there is less total sulphated glycosaminoglycan available to interact with this tryptase.

Protein modelling, crystallographic and site-directed mutagenesis studies have revealed that the GAG-binding domain on the surface of each mast cell granule protease is conformation dependent. Because the granule storage capacity of a mast cell is limited, serglycin proteoglycans appear to be used by this cell to ensure that only properly folded proteases are targeted to the secretory granule. Mast cells must be able to respond quickly to changes in the immune status of their microenvironments⁹. Thus, the proteases must be packaged in granules in their mature forms. A larger number of distinct glycosyltransferases and sulphotransferases are required to synthesize heparin and chondroitin sulphate. Why mast cells place very different types of negatively charged GAGs onto serglycin has not been deduced. Nevertheless, the observation that the post-translational modification of serglycin is conserved in evolution from mice to humans indicates that it is very important to the function of the mast cell. The observation that the mBMMCs developed with IL-3 from NDST-2-null mice can still package tryptases (Fig. 4) that can cleave tosyl-Gly-Pro-Arg-pNA is consistent with the *in vivo* data (Fig. 2) that showed that fully sulphated heparin is not essential for tryptase expression. Because the chondroitin-sulphate-rich mast cells residing in the jejunum contain substantial amounts of mMCP-2 (Fig. 3), heparin is also not essential for expression of this chymase.

One of the mysteries in immunology is how mast cells can store large amounts of enzymatically active proteases in their granules in such a way that these enzymes do not undergo rapid autolysis. Based on data from numerous *in vitro* studies it has been speculated that heparin is an essential co-factor in mast cells for the dipeptidyl-peptidase-I-mediated conversion of zymogens into active proteases²⁰. Whether NDST-2-null mBMMCs were cultured in the absence (Fig. 4) or presence (Fig. 5) of fibroblasts, the resulting cells could not store mMCP-5 and mMC-CPA in their granules, even though they contained high levels of the chymase and exopeptidase transcripts. mMCP-5 and mMC-CPA have relative molecular masses of 30,000 and 36,000 (*M_r* = 30K and 36K), respectively. Consistent with the SDS-PAGE/immunoblot data, the levels of the 30K and 36K proteins in wild-type mBMMCs was greater than that in NDST-2-null mBMMCs (Fig. 5b). These data indicate that fully sulphated heparin is essential for the expression of two mast-cell chymases and an exopeptidase. The mast cell uses transcriptional and post-transcriptional mechanisms to alter the levels of its individual protease transcripts. This is the first example where the expression of one granule constituent markedly affects the expression of another at the post-translational level. Thus, the mast cell uses multiple mechanisms to regulate the levels of its various granule proteases. □

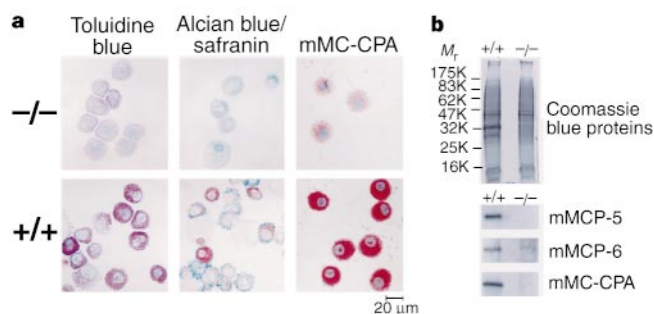


Figure 5 Histochemistry, immunohistochemistry and SDS-PAGE/immunoblot analysis of fibroblast/mBMMC co-cultures. **a**, Co-cultured mBMMCs, developed from NDST-2-null mice and their wild-type littermates, were evaluated histochemically for their proteoglycan content and immunohistochemically for their mMCP-CPA content. **b**, SDS-PAGE/immunoblot analysis was used to evaluate the relative levels of Coomassie blue stainable proteins, as well as mMCP-5, mMCP-6 and mMC-CPA proteins in the two populations of co-cultured mast cells.

Methods

Generation of NDST-2-null mice. We screened a 129/Sv mouse genomic library with a probe corresponding to residues 1488–2128 in the transcript² to isolate the coding portion of the NDST-2 gene. We prepared a targeting construct in which amino-acids 539 to 580 of the mouse NDST-2 gene were replaced by the Neomycin resistance (Neo^r) gene. The carboxy-terminal halves of the varied NDSTs exhibit considerable homology with other sulphotransferases and previous functional studies carried out on recombinant NDSTs have revealed that the *N*-sulphotransferase activity of this family of enzymes is entirely contained within residues 479–880. Thus, to ensure a nonfunctional enzyme, we also placed a premature stop codon after residue 538. Embryonic stem cells were transfected with the targeting construct. After positive selection, a clone that possessed a disrupted NDST-2 allele was injected into blastocysts to eventually obtain NDST-2-null mice.

Helminth infection and IgE-dependent passive anaphylaxis. To evaluate the consequence of disrupted heparin expression on the chondroitin sulphate population of mast cells in the jejunum of helminth-infected mice, NDST-2-null mice and their wild-type littermates were each infected orally with 400 *T. spiralis* larvae, as described⁹. We examined the jejunal mast cells at the height of the mastocytosis at day 14. For the IgE-dependent PCA reactions, PBS (20 µl) with or without 20 ng of anti-dinitrophenyl IgE (Sigma) was injected into separate ears of each NDST-2-null and wild-type mouse. Twenty-four hours later, a 100-µl solution of 0.5% Evans blue containing 100 µg dinitrophenyl albumin was injected into the tail vein of each animal; the uptake of the dye into the two ears was then compared ~30 min later.

Histochemistry, enzyme cytochemistry and immunohistochemistry. mBMMCs were obtained as described¹⁶ by culturing isolated progenitors from NDST-2-null mice and their wild-type littermates for 3 weeks in IL-3-supplemented medium. Some of these mBMMCs were then co-cultured with 3T3 fibroblasts¹⁷. For histological examination, we stained cytocentrifuge preparations of mBMMCs and 1.5-µm-thick glycolmethacrylate sections of various tissues with a 0.5% solution of toluidine blue in 0.6 M HCl. mBMMCs were also stained with 0.5% alcian blue in 0.3% acetic acid and then counterstained with 0.1% safranin O in 1% acetic acid. Those mast cells that contained esterase activities were identified using a modification⁹ of the method of ref. 21. We incubated tissue sections and cultured mast cells with antibodies specific for mMCP-2 (ref. 22), mMCP-4 (R.L.S., unpublished data), mMCP-5 (ref. 23), mMCP-6 (ref. 24), mMCP-7 (ref. 24) and mMCP-CPA²⁵. To confirm the immunocytochemical data, lysates of pooled mBMMCs developed from NDST-2-null mice and their wild-type littermates were electrophoresed in 12% SDS-PAGE gels. Protein blots prepared from the separated proteins were probed with anti-mMCP-5, anti-mMCP-6 or anti-mMCP-CPA antibodies.

RNA blot analysis. We isolated total RNA¹⁵ from the ear and tongue of NDST-2-null mice and their wild-type littermates, as well as the mBMMCs developed from these mice. RNA blots were prepared and analysed with NDST-2 (the 641-base-pair fragment that corresponds to residues 1488–2128 in the transcript²), NDST-1 (ref. 4), mMCP-5 (ref. 12), mMCP-6 (ref. 11), mMCP-7 (ref. 13), mMCP-CPA or β-actin²⁶ radiolabelled cDNAs.

N-Sulphotransferase, proteoglycan, tryptase and histamine biochemical assays. We evaluated mBMMCs for their *N*-sulphotransferase activities as described²⁷. To evaluate the consequences of disrupting the NDST-2 gene on proteoglycan expression in mast cells, mBMMCs from NDST-2-null mice and their wild-type littermates were co-cultured with fibroblasts for an additional 2–4 weeks. The adherent cells in the resulting co-cultures were radiolabelled in medium containing 150 µCi [³⁵S]sodium sulphate¹⁶. After this step, each culture was washed and then cultured for an additional 2 h in normal medium. The mast cell/fibroblast co-cultures were then exposed to a solution containing 0.25% trypsin and 1 mM EDTA for 5 min to remove the remaining fibroblast-derived [³⁵S]proteoglycans. After the detached fibroblasts and mast cells were centrifuged, the cell pellets were disrupted and the liberated [³⁵S]proteoglycans were subjected to CsCl density-gradient centrifugation²⁸. The purified proteoglycans were then treated with either chondroitinase ABC²⁹ to degrade their chondroitin-sulphate chains or nitrous acid²⁸ to degrade their heparin-like chains. The intracellular, heparin-like population of GAGs produced by the NDST-2-null and wild-type mBMMCs were subjected to DEAE-cellulose column chromatography to evaluate their overall net charges. In one

experiment, samples also were subjected to anti-thrombin-III affinity chromatography to evaluate the presence of a specific structural motif⁸ within the chain.

To determine whether or not the immunoreactive tryptases present in the mBMMCs developed from NDST-2-null mice were enzymatically active, we tested the ability of cell lysates to cleave tosyl-Gly-Pro-Arg-pNA³⁰. Finally, an ELISA kit (ICN Pharmaceuticals) was used to measure the levels of histamine in sonicates of mBMMCs and peritoneal cavity exudates.

Received 15 March; accepted 16 June 1999.

- Orellana, A., Hirschberg, C. B., Wei, Z., Swidler, S. J. & Ishihara, M. Molecular cloning and expression of glycosaminoglycan N-acetylglucosaminyl N-deacetylase/N-sulphotransferase from a heparin-producing cell line. *J. Biol. Chem.* **269**, 2270–2276 (1994).
- Eriksson, I., Sandback, D., Ek, B., Lindahl, U. & Kjellén, I. cDNA cloning and sequencing of mouse mastocytoma glucosaminyl N-deacetylase/N-sulphotransferase, an enzyme involved in the biosynthesis of heparin. *J. Biol. Chem.* **269**, 10439–10443 (1994).
- Lindahl, U. & Kjellén, I. in *Biology of Proteoglycans* (eds Land, D. A. & Lindahl, U.) 159–189 (Edward Arnold, London, 1989).
- Hashimoto, Y., Orellana, A., Gil, G. & Hirschberg, C. B. Molecular cloning and expression of rat liver N-heparan sulfate sulphotransferase. *J. Biol. Chem.* **267**, 16744–16750 (1992).
- Aikawa, J. & Esko, J. D. Molecular cloning and expression of a third member of the heparan sulfate/heparin GlcNAc N-deacetylase/N-sulphotransferase family. *J. Biol. Chem.* **274**, 2690–2695 (1999).
- Stevens, R. L. *et al.* Intestinal mucosa mast cells from rats infected with *Nippostrongylus brasiliensis* contain protease-resistant chondroitin sulfate di-B proteoglycans. *J. Immunol.* **137**, 291–295 (1986).
- Le Trong, H. *et al.* Amino acid sequence of a mouse mucosal mast cell protease. *Biochemistry* **28**, 391–395 (1989).
- Serafin, W. E. *et al.* Identification and molecular cloning of a novel mouse mucosal mast cell serine protease. *J. Biol. Chem.* **265**, 423–429 (1990).
- Friend, D. S. *et al.* Mast cells that reside at different locations in the jejunum of mice infected with *Trichinella spiralis* exhibit sequential changes in their granule ultrastructure and chymase phenotype. *J. Cell Biol.* **135**, 279–290 (1996).
- Serafin, W. E. *et al.* Cloning of the cDNA and gene for mouse mast cell protease-4. *J. Biol. Chem.* **266**, 1934–1941 (1991).
- Reynolds, D. S., Gurley, D. S., Austen, K. F. & Serafin, W. E. Cloning of the cDNA and gene of mouse mast cell protease-6. *J. Biol. Chem.* **266**, 3847–3853 (1991).
- McNeil, H. P., Austen, K. F., Somerville, L. L., Gurish, M. F. & Stevens, R. L. Molecular cloning of the mouse mast cell protease-5 gene. *J. Biol. Chem.* **266**, 20316–20322 (1991).
- McNeil, H. P. *et al.* Isolation, characterization, and transcription of the gene encoding mouse mast cell protease 7. *Proc. Natl Acad. Sci. USA* **89**, 11174–11178 (1992).
- Lützelshwab, C., Huang, M. R., Kullberg, M. C., Aveskog, M. & Hellman, L. Characterization of mouse mast cell proteases-8, the first member of a novel subfamily of mouse mast cell serine proteases, distinct from both the classical chymases and tryptases. *Eur. J. Immunol.* **28**, 1022–1033 (1998).
- Stevens, R. L. *et al.* Strain-specific and tissue-specific expression of mouse mast cell secretory granule proteases. *Proc. Natl Acad. Sci. USA* **91**, 128–132 (1994).
- Razin, E., Stevens, R. L., Akiyama, F., Schmid, K. & Austen, K. F. Culture from mouse bone marrow of a subclass of mast cells possessing a distinct chondroitin sulfate proteoglycan with glycosaminoglycans rich in N-acetylgalactosamine-4,6-disulfate. *J. Biol. Chem.* **257**, 7229–7236 (1982).
- Levi-Schaffer, F., Austen, K. F., Gravalles, P. M. & Stevens, R. L. Coculture of interleukin 3-dependent mouse mast cells with fibroblasts results in a phenotypic change of the mast cells. *Proc. Natl Acad. Sci. USA* **83**, 6485–6488 (1986).
- Lindahl, U., Backstrom, G., Thunberg, L. & Leder, I. G. Evidence for a 3-O-sulfated D-glucosamine residue in the antithrombin-binding sequence of heparin. *Proc. Natl Acad. Sci. USA* **77**, 6551–6555 (1980).
- Sakai, K., Ren, S. & Schwartz, L. B. A novel heparin-dependent processing pathway for human tryptase, autocatalysis followed by activation with dipeptidyl peptidase I. *J. Clin. Invest.* **97**, 988–995 (1996).
- McGuire, M., Lipsky, P. E. & Thiele, D. L. Generation of active myeloid and lymphoid granule serine proteases requires processing by the granule thiol protease dipeptidyl peptidase I. *J. Biol. Chem.* **268**, 2458–2467 (1993).
- Leder, I. D. The chloroacetate ester reaction. A useful means of histological diagnosis of hematological disorders from paraffin sections of skin. *Am. J. Dermatopathol.* **1**, 39–42 (1979).
- Ghildyal, N., Friend, D. S., Nicodemus, C. F., Austen, K. F. & Stevens, R. L. Reversible expression of mouse mast cell protease 2 mRNA and protein in cultured mast cells exposed to interleukin 10. *J. Immunol.* **151**, 3206–3214 (1993).
- McNeil, H., Frenkel, D. P., Austen, K. F., Friend, D. S. & Stevens, R. L. Translation and granule localization of mouse mast cell protease-5: immunodetection with specific antipeptide Ig. *J. Immunol.* **149**, 2466–2472 (1992).
- Ghildyal, N. *et al.* Fate of two mast cell tryptases in V3 mastocytosis and normal BALB/c mice undergoing passive systemic anaphylaxis. Prolonged retention of exocytosed mMCP-6 in connective tissues and rapid accumulation of enzymatically active mMCP-7 in the blood. *J. Exp. Med.* **184**, 1061–1073 (1996).
- Gurish, M. F. *et al.* Mouse mast cells that possess segmented/multi-lobular nuclei. *Blood* **90**, 382–390 (1997).
- Spiegelman, B. M., Frank, M. & Green, H. Molecular cloning of mRNA from 3T3 adipocytes. Regulation of mRNA content for glycerophosphate dehydrogenase and other differentiation-dependent proteins during adipocyte development. *J. Biol. Chem.* **258**, 10083–10089 (1983).
- Humphries, D. E., Sullivan, B. M., Aleixo, M. D. & Stow, J. L. Localization of human heparan glucosaminyl N-deacetylase/N-sulphotransferase to the trans-Golgi network. *Biochem. J.* **325**, 351–357 (1997).
- Stevens, R. L. & Austen, K. F. Effect of p-nitrophenyl-β-D-xyloside on proteoglycan and glycosaminoglycan biosynthesis in rat serosal mast cell cultures. *J. Biol. Chem.* **257**, 253–259 (1982).
- Saito, H., Yamagata, T. & Suzuki, S. Enzymatic methods for the determination of small quantities of isomeric chondroitin sulfates. *J. Biol. Chem.* **243**, 1536–1542 (1968).
- Huang, C. *et al.* Induction of a selective and persistent extravasation of neutrophils into the peritoneal cavity by the tryptase mouse mast cell protease 6. *J. Immunol.* **160**, 1910–1919 (1998).

Acknowledgements. This work was supported by the US VA and NIH. The technical assistance of X. Hu is acknowledged.

Correspondence and requests for materials should be addressed to R.L.S. (e-mail: rstevens@rics.bwh.harvard.edu) or D.E.H. (e-mail: Humphries.Donald_E@BOSTON.VA.GOV).

Abnormal mast cells in mice deficient in a heparin-synthesizing enzyme

Erik Forsberg^{*†}, Gunnar Pejler^{†‡}, Maria Ringvall^{*}, Carolina Lunderius^{*}, Bianca Tomasini-Johansson[‡], Marion Kusche-Gullberg[§], Inger Eriksson[‡], Johan Ledin[‡], Lars Hellman^{*} & Lena Kjellén[‡]

Department of ^{*}Cell and Molecular Biology, Box 596, Department of [§]Medical Biochemistry and Microbiology, Box 582, University of Uppsala, S-751 23 Uppsala, Sweden

[‡]Department of Veterinary Medical Chemistry, Swedish University of Agricultural Sciences, Box 575, S-75123 Uppsala, Sweden

[†] These authors contributed equally to this work

Heparin is a sulphated polysaccharide, synthesized exclusively by connective-tissue-type mast cells¹ and stored in the secretory granules in complex with histamine and various mast-cell proteases². Although heparin has long been used as an anti-thrombotic drug, endogenous heparin is not present in the blood, so it cannot have a physiological role in regulating blood coagulation. The biosynthesis of heparin involves a series of

enzymatic reactions, including sulphation at various positions^{1,3}. The initial modification step, catalysed by the enzyme glucosaminyl *N*-deacetylase/*N*-sulphotransferase-2, NDST-2 (refs 4–7), is essential for the subsequent reactions. Here we report that mice carrying a targeted disruption of the gene encoding NDST-2 are unable to synthesize sulphated heparin. These NDST-2-deficient mice are viable and fertile but have fewer connective-tissue-type

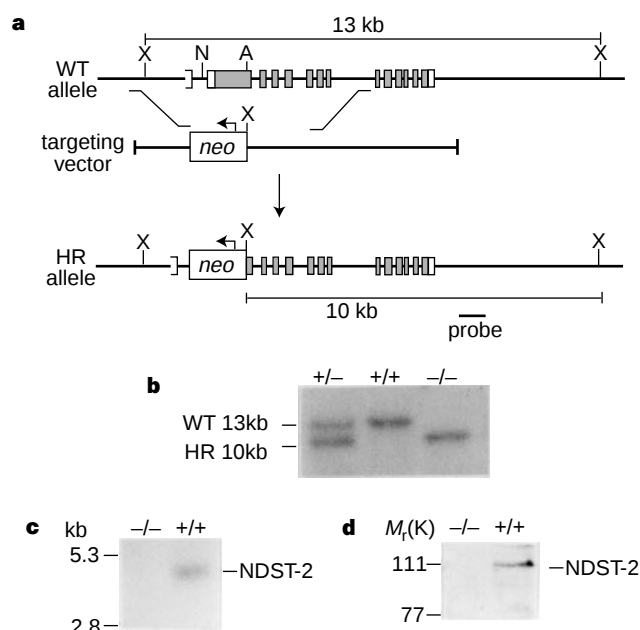


Figure 1 Targeted disruption of the *NDST-2* locus and establishment of a mutant mouse strain. **a**, The *NDST-2* gene (WT, wild type; top), the targeting vector (middle) and the disrupted *NDST-2* locus (HR, homologous recombination; bottom). A 1,329-bp deletion, including 307 bp of 5'-untranslated and 780 bp of coding sequence, was replaced with a *pgk-neo* cassette. The expected size of an *Xba*I digestion product of the gene, hybridizing with the indicated probe, is shown for the wild-type locus (top) and for the mutated allele (bottom). Restriction enzymes: A, *Apa*I; N, *Nhe*I; X, *Xba*I. **b**, Genomic DNA isolated from F₂ litter mates of the intercross of *NDST-2*^{+/-} heterozygous mice was digested with *Xba*I, blotted, and hybridized with the probe in **a**. Genotypes of the progeny are indicated at the top of each lane. **c**, Northern-blot analysis of mRNA from testes of *NDST-2*^{-/-} (left lane) and *NDST-2*^{+/-} mice (right lane). The probe corresponds to the fragment between the *Nhe*I/*Apa*I sites in **a**. **d**, Western-blot analysis of peritoneal cell extracts from *NDST-2*^{-/-} and *NDST-2*^{+/-} mice. Binding of the peptide antibody recognizing NDST-2 (ref. 6, peptide 2) was detected by using electrochemiluminescence (ECL) (Amersham). The 110 kD NDST-2 protein is seen in extracts from wild-type mice but not from *NDST-2*-null mice.

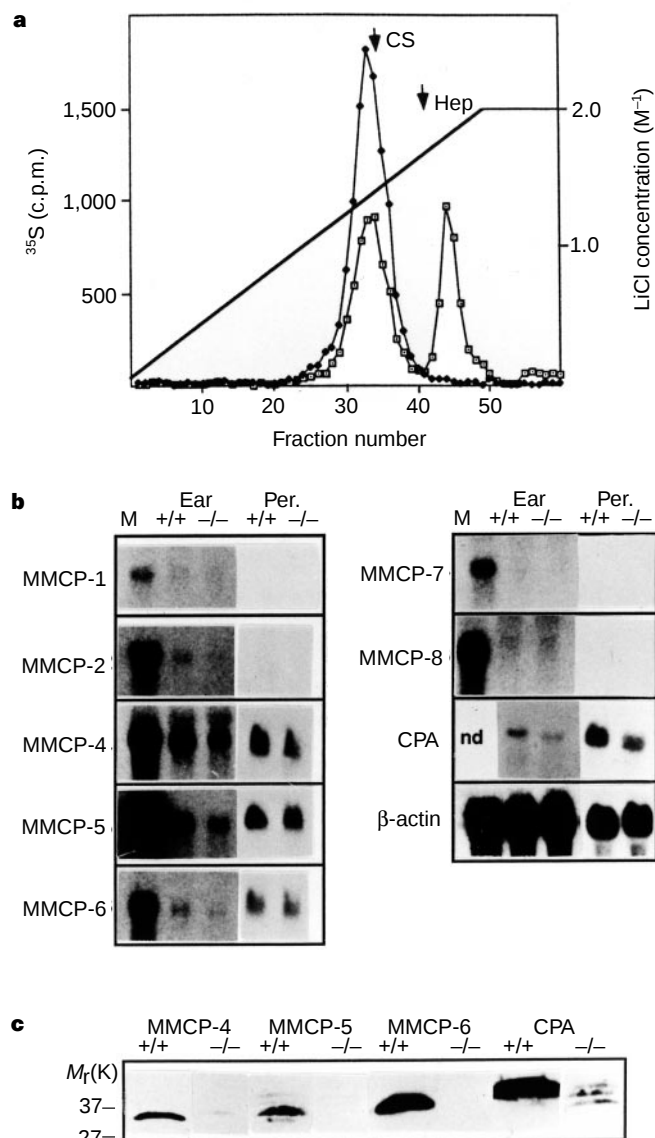


Figure 2 Absence of normally sulphated heparin and mast-cell proteases in peritoneal cells from *NDST-2*^{-/-} mice. **a**, [³⁵S]macromolecules from *NDST-2*^{+/-} (open symbols) and *NDST-2*^{-/-} mice (filled symbols), respectively, were mixed with unlabelled glycosaminoglycan standards (chondroitin sulphate and heparin) and applied to a 1-ml DEAE-Sephacel column that was eluted with a linear gradient of LiCl in 0.05 M acetate buffer, pH 4.0. Elution positions of chondroitin sulphate (CS) and heparin (Hep) are indicated. **b**, Northern blot under high-stringency conditions²⁵ of total RNA extracted from peritoneal cells (Per.) and ear tissue from *NDST-2*^{+/-} and *NDST-2*^{-/-} mice and from mouse mastocytoma tissue (M). RNA was analysed for transcripts encoding MMCP-1, -2, -4, -5, -6, -7 and -8, and carboxypeptidase A (CPA). Probes were specific for the various mast-cell proteases²⁵. β -Actin was used as a loading control. **c**, Western-blot analysis of mast-cell proteases using specific antisera against MMCP-2, -4, -5, -6 and -7 and carboxypeptidase A. Peritoneal-cell extracts from *NDST-2*^{+/-} and *NDST-2*^{-/-} mice were subjected to SDS-PAGE and western blotted. Filters were developed by using the ECL system after incubation with anti-rat or anti-rabbit immunoglobulin conjugated with horseradish peroxidase.

mast cells; these cells have an altered morphology and contain severely reduced amounts of histamine and mast-cell proteases. Our results indicate that one site of physiological action for heparin could be inside connective-tissue-type mast cells, where its absence results in severe defects in the secretory granules.

Two NDSTs, NDST-1 and NDST-2, have been described¹⁵⁻⁹, and the cloning of a third isoform was recently reported¹⁰. Connective-tissue-type mast cells express little or no NDST-1, but contain large amounts of the transcript encoding NDST-2 (ref. 7). To establish a mouse strain with defective synthesis of heparin, we introduced a mutation into the *NDST-2* gene. Using gene targeting in embryonic stem cells, the sequence encoding the cytoplasmic and *trans*-Golgi membrane domain, as well as part of the luminal, catalytically active domains of the enzyme, was deleted by inserting a neomycin-resistance cassette (*neo*; Fig. 1a). Homologous recombination in the *NDST-2* locus was found in two clones, which, after hybridization with a *neo* probe, showed no additional sites of integration (results not shown). Germline transmission was achieved with both clones. Heterozygous mice did not display any obvious abnormalities compared to their littermates.

We analysed genotypes of 167 pups (Fig. 1b) from intercrosses between heterozygous mice, which showed the expected mendelian hereditance (25.1% wild type, 52.1% heterozygous, 22.8% homozygous), indicating that NDST-2 is not crucial for embryonic development. Northern-blot analysis of testis messenger RNA (Fig. 1c) and polymerase chain reaction with reverse transcription (RT-PCR) of testis and liver RNAs (results not shown) showed that the replaced part of the gene is not expressed in homozygous mice. Furthermore, the NDST-2 protein is present in extracts of peritoneal cells of *NDST-2*^{+/+} mice, as shown by western blotting (Fig. 1d), whereas no NDST-2 could be detected in cells isolated from *NDST-2*^{-/-} mice, indicating that this was a true null mutation. Female and male *NDST-2*^{-/-} mice were fertile and showed no obvious pathological phenotype at 20 months. In addition, histological examination of liver, kidneys, testis, lung, heart, skeletal muscle and spleen revealed no apparent defects.

To test whether *NDST-2*^{-/-} cells could synthesize heparin, peri-

toneal cells that are a mixture of macrophages, lymphocytes and mast cells were labelled with ³⁵S-sulphate, followed by purification and characterization of ³⁵S-labelled macromolecules. Equal amounts of [³⁵S]chondroitin sulphate, which is synthesized by both lymphocytes and macrophages, were recovered from *NDST-2*^{+/+} and *NDST-2*^{-/-} cells, whereas the amount of [³⁵S]heparin and/or heparan [³⁵S]sulphate was greatly reduced in *NDST-2*^{-/-} cells (Table 1). Results were similar in four experiments. To distinguish between heparin and heparan sulphate, [³⁵S]polysaccharide chains were released from their core proteins by alkali and analysed by DEAE ion-exchange chromatography, when heparin elutes later in the salt gradient owing to its higher charge density. None of the mutant cell [³⁵S]macromolecules eluted at the position of the heparin standard (Fig. 2a), indicating that ³⁵S-labelled heparin was absent from *NDST-2*^{-/-} cells. Preliminary results using ³H-glucosamine-labelled peritoneal cells indicate that the unsulphated heparin precursor is formed in *NDST-2*^{-/-} mast cells but in about half the normal amount, which corresponds roughly to the observed reduction in the number of mast cells (Table 1). The lack of ³⁵S-sulphated heparin in NDST-2-deficient cells indicates that this enzyme is crucial for heparin biosynthesis. The role of NDST-2 in the biosynthesis of heparan sulphate is less obvious: structural analysis of liver heparan sulphate (J.L., I.E., E.F. and L.K., unpublished results) reveals no differences in the *N*-sulphation patterns of polysaccharides isolated from control and *NDST-2*^{-/-} mice—58% of the glucosamine residues in both preparations carry *N*-sulphated groups (data not shown), so NDST-1 and -3 might compensate for the lack of NDST-2 in the biosynthesis of heparan sulphate. Proteoglycans carrying heparan sulphate have been implicated in several physiological processes, including organogenesis, during embryonic development¹¹⁻¹⁷. Our results, demonstrating a lack of influence of NDST-2 on embryonic development and in adult life, indicate that the other NDST isoforms may be more important in the biosynthesis of heparan sulphate, an idea supported by studies of mice with a targeted mutation in *NDST-1* (M.R., J.L., I.E., L.K. and E.F., unpublished results).

The degree of maturation and differentiation of mast cells is

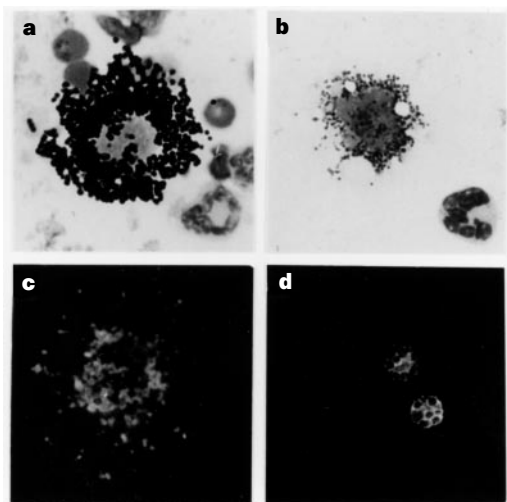


Figure 3 Absence of normal peritoneal mast cells in *NDST-2*^{-/-} mice. **a-d**, Cytopins of peritoneal cells from *NDST-2*^{+/+} mice (**a, c**) and *NDST-2*^{-/-} mice (**b, d**) were stained with May-Grünwald, followed by 5% Giemsa stain in PBS (**a, b**), or fixed in ethanol and allowed to bind polyclonal rabbit antibodies against histamine (Euro-Diagnostica), followed by incubation with Cy3-labelled goat anti-rabbit immunoglobulin (Jackson ImmunoResearch) (**c, d**). The presence of neutrophils in **a** and **b** is caused by intraperitoneal injection of 2 µg lipopolysaccharide (LPS) 8 h before the peritoneal lavage. Neutrophil numbers were comparable in LPS-injected *NDST-2*^{+/+} and *NDST-2*^{-/-} mice (15–25% (*n* = 7 for both control and mutant mice)).

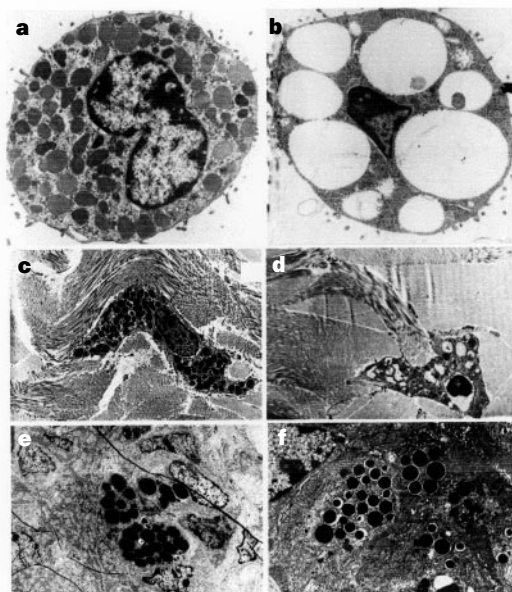


Figure 4 Altered morphology of heparin-containing mast cells. **a-f**, Transmission electron micrographs of mast cells from the peritoneum (**a, b**), from the skin of the back (**c, d**), and from the jejunum of the intestine (**e, f**) of *NDST-2*^{+/+} mice (**a, c, e**) and *NDST-2*^{-/-} mice (**b, d, f**).

Table 1 Properties of *NDST-2*^{-/-} and *NDST-2*^{+/-} peritoneal cells

Source of cells	³⁵ S-chondroitin sulphate (c.p.m. × 10 ⁻³ /10 ⁶ cells)*	³⁵ S-heparan sulphate/heparin (c.p.m. × 10 ⁻³ /10 ⁶ cells)†	Mast cells (%)‡	Trypsase activity (units)§	Chymase activity (units)§	Histamine content (nmol/10 ⁶ cells)¶
<i>NDST-2</i> ^{+/-} mice	4.9	8.3	2.0 ± 0.36 (n = 3)	1.0 ± 0.91 (n = 11)	0.67 ± 0.52 (n = 11)	1.7 ± 0.22 (n = 4)
<i>NDST-2</i> ^{-/-} mice	4.9	1.0	0.56 ± 0.19 (n = 3)	0.027 ± 0.031 (n = 11)	0.016 ± 0.023 (n = 11)	0.11 ± 0.044 (n = 4)

* [³⁵S]macromolecules sensitive to chondroitinase ABC digestion.

† [³⁵S]macromolecules sensitive to treatment with HNO₂ at pH 1.5.

‡ Mast cells were detected and counted after combined stainings with May-Grünwald/Giemsa and fluorescent antibodies against histamine (Fig. 3 legend).

§ Enzyme activities were assayed using chromogenic peptide substrates: trypsin, H-D-Ile-Pro-Arg-pNA (S-2288; Chromogenix); chymase, MeO-Suc-Arg-Pro-Tyr-pNA (S-2586, Chromogenix). One unit of trypsin/chymase activity corresponds to the activity that yields an optical absorbance change at 405 nm of 1.0 per hour for 10⁶ cells.

¶ Histamine was quantified by enzyme immunoassay (Immunotech). Results are shown with standard deviations.

intimately linked to the repertoire of proteases they express^{18,19}: for example, mature connective-tissue mast cells express the chymases known as mouse mast-cell protease-4 (MMCP-4), -5 and -2, the trypsinases MMCP-6 and -7, and carboxypeptidase A, whereas mast cells of the mucosal type preferentially express MMCP-1 and -2. Measurement of mast-cell protease activity in peritoneal cell extracts from *NDST-2*^{-/-} mice revealed barely any chymase or trypsin activity (Table 1). Western blot analysis showed that MMCP-4, -5 and -6 and carboxypeptidase A antigens were present in peritoneal cells from *NDST-2*^{+/-} mice (Fig. 2c; MMCP-2 and -7 were not detectable). In contrast, peritoneal cell extracts from *NDST-2*^{-/-} mice were virtually devoid of protease antigens (Fig. 2c). MMCP-2, -4, -5 and -6 and carboxypeptidase were expressed in both *NDST-2*^{+/-} and *NDST-2*^{-/-} animals, as indicated by northern blots of total RNA extracted from ears and peritoneal cells (Fig. 2b). Signals were generally weaker in the *NDST-2*^{-/-} mice, in agreement with the observed reduction in the number of mast cells (Table 1). Thus, the absence of heparin does not affect the levels of mast-cell protease mRNA but it does affect the amount of protease stored in the secretory granules. Presumably, the high negative charge of normally sulphated heparin is required for the assembly of positively charged mast-cell mediators, including the mast-cell proteases and histamine. Further, as connective-tissue-type mast cells from both *NDST-2*^{+/-} and *NDST-2*^{-/-} mice express the same repertoire of proteases, the absence of heparin proteoglycan does not affect the degree of mast-cell maturation.

To find out the effect of a lack of heparin on the mast-cell phenotype, we stained peritoneal cells with toluidine blue (0.1% in PBS buffer) to quantify the number of mast cells present. The cell population from control mice showed the normal frequency of connective-tissue-type mast cells (2%), but we detected no toluidine-blue-staining in homozygous littermates (data not shown). May-Grünwald/Giemsa staining revealed *NDST-2*^{-/-} mast cells with a reduced number of weakly staining, smaller granules and large empty vacuoles (Fig. 3b). The *NDST-2*^{+/-} mast cells, by contrast, contained numerous densely stained granules (Fig. 3a). Histamine-positive cells were detected in the *NDST-2*^{-/-} peritoneal cell population by immunostaining (Fig. 3d): these cells had a raspberry-like appearance, with most of the histamine being present between the large vacuoles, probably in the small granular structures observed after May-Grünwald and Giemsa staining (Fig. 3b).

To test whether *NDST-2*^{-/-} mast cells could still initiate an inflammatory response, we injected mice intraperitoneally with IgE and then with antibodies against IgE, and studied the neutrophil influx. We found no obvious difference in the recruitment of neutrophils in the peritoneum of *NDST-2*^{+/-} and *NDST-2*^{-/-} mice at 3 and 7 h after injection (*NDST-2*^{+/-}: 38.2 ± 13.7% neutrophils at 3 h (n = 8) and 39.8 ± 8.2% at 7 h (n = 8); *NDST-2*^{-/-}: 31.9 ± 14.7% neutrophils at 3 h (n = 8) and 36.1 ± 14.0% at 7 h (n = 8)). *In vitro*, *NDST-2*^{-/-} cells responded specifically to the addition of IgE and anti-IgE by releasing a substantial amount of their stored histamine (27 ± 7.8 pmol per 10⁶ *NDST-2*^{-/-} peritoneal cells (n = 4), compared to 56 ± 8.1 pmol per 10⁶ *NDST-2*^{+/-} peritoneal cells (n = 4). As the number of mast cells is reduced in the

mutant mice (Table 1), the amount of histamine released per peritoneal mast cell is actually similar or even larger in *NDST-2*^{-/-} as compared to *NDST-2*^{+/-} mice, despite the marked overall decrease in histamine in *NDST-2*^{-/-} mast cells (Table 1).

Examination of the peritoneal cells using transmission electron microscopy revealed striking differences between mast cells from *NDST-2*^{+/-} mice (Fig. 4a) and *NDST-2*^{-/-} mice (Fig. 4b): *NDST-2*^{+/-} mast cells contained numerous electron-dense granules, whereas *NDST-2*^{-/-} mast cells have only a few granules and apparently 'empty' vacuoles. Mast cells from skin tissue sections from *NDST-2*^{-/-} mice had similar abnormalities (Fig. 4d) (skin mast cells from *NDST-2*^{+/-} mice were normal; Fig. 4c). May-Grünwald/Giemsa staining and histamine immunostaining showed that the number of mast cells in skin from the back of *NDST-2*^{-/-} mice was reduced by about half. Skin mast cells contain heparin, whereas intestinal mucosa mast cells contain chondroitin sulphate²⁰. As *NDST* enzymes are not involved in the biosynthesis of this latter polysaccharide, it was important to verify that mucosal mast cells were unaffected by the absence of the *NDST-2* gene. Transmission electron microscopy on sections of the jejunum of the intestine confirmed the presence of mast cells with electron-dense granular structures in both *NDST-2*^{+/-} (Fig. 4e) and *NDST-2*^{-/-} mice (Fig. 4f).

Mast cells effect allergic reactions and participate in other acute and chronic inflammatory conditions²¹. Our *NDST-2*-deficient mice provide a model for studying the role of heparin in mast-cell function *in vivo* and in the development of mast cells, as well as its effect on their mediator content. □

Methods

Construction of the targeting vector. A 15-kb genomic clone was isolated from a 129/Sv genomic library (Stratagene) and characterized⁷. A 2.8-kb *EcoRI*/*ApaI* fragment was subcloned into a pBSIISK+ vector (Stratagene). Subsequently, a *pgk-neo* vector was cloned into the *ApaI* and an *NheI* site in the opposite direction relative to the transcription of the *NDST-2* gene. This ligation generated a 1.5-kb arm of homology fused to the *neo* cassette. Finally, using *ApaI*, a 6-kb arm of homology was subcloned from the genomic clone into the *ApaI* site of the shortarm *neo* vector.

ES cells and generation of mutant mice. Embryonic stem (ES) cells were cultured as described²². The targeting vector, linearized with *NotI*, was electroporated into R1 ES cells²³. Clones that had survived selection in 350 µg ml⁻¹ G418 (Gibco BRL) were Southern blotted using an 800-bp *SphI*/*ApaI* fragment, located 3' of the targeting vector, as probe. The two clones that had undergone homologous recombination were injected into C57BL/6 blastocysts. Chimaeric male founder mice were crossed with C57BL/6 females to obtain heterozygous F₁ offspring. Tail biopsies were genotyped as described²². By intercrossing heterozygous mice, an *NDST-2*-deficient mouse strain was established and the phenotype was analysed using F₂ littermates. Crossing of the chimaeric male founder mice with 129/SvJ female mice confirmed that the phenotype did not differ in an inbred 129/SvJ genetic background (results not shown).

Isolation of peritoneal cells. Peritoneal cells were obtained by peritoneal lavage with 10 ml cold PBS. Cells were centrifuged at 300g and washed once in PBS before use.

Glycosaminoglycan analysis. Peritoneal cells from two control and two mutant mice, respectively (4.5–7 × 10⁶ cells), were labelled *in vitro* with

0.5 mCi of carrier-free ^{35}S -sulphate by overnight incubation at 37°C . ^{35}S -labelled macromolecules, obtained after solubilization of cells with buffer containing 1% Triton X-100, were purified by DEAE ion-exchange chromatography essentially as described²⁴. Samples were desalted before treatment with chondroitinase ABC (which degrades chondroitin sulphate) or HNO_2 at pH 1.5 (degrades heparan sulphate and heparin), followed by gel chromatography on Sephadex G-50 (ref. 24).

Antisera. Antisera against MMCP-5, -6 and -7 and carboxypeptidase A were raised in rats against fusion proteins derived from the corresponding full-length clones²⁵; antisera against MMCP-2 and -4 were raised in rabbits against non-homologous peptide sequences conjugated with keyhole limpet haemocyanin.

IgE/anti-IgE challenge. For studies of neutrophil recruitment, mice were injected with $1\ \mu\text{g}$ of murine IgE (PharMingen), then with $1\ \mu\text{g}$ rat anti-murine IgE (PharMingen) 20 min later. Peritoneal cells were collected 3 and 7 h after IgE injection. Cytospins were stained with May-Grünwald stain then with 5% Giemsa stain in PBS and the percentage of neutrophils in the peritoneal cell population was determined. For *in vitro* studies of histamine release, 6×10^5 peritoneal cells in $300\ \mu\text{l}$ M2 medium (Sigma) were incubated at 37°C with $0.25\ \mu\text{g}$ IgE for 12 min, followed by a 12-min incubation with $2.5\ \mu\text{g}$ of anti-IgE. After centrifugation, histamine content in the supernatant and in the residual cells was quantified by EIA (Immunotech). The amounts of histamine released without IgE/anti IgE addition were $27 \pm 3.7\ \text{pmol per } 10^6\ \text{NDST-2}^{+/+}$ cells ($n = 4$) and $5.7 \pm 2.0\ \text{pmol per } 10^6\ \text{NDST-2}^{-/-}$ cells ($n = 4$).

Transmission electron microscopy. Cells and tissues were fixed in 2% glutaraldehyde, incubated in 1% OsO_4 /PBS, dehydrated and embedded in TAAB-B12 resin. Sections were analysed at 60 kV in a Philips CM10 microscope.

Received 6 May; accepted 2 July 1999.

- Kjellén, L. & Lindahl, U. Proteoglycans: Structure and interactions. *Annu. Rev. Biochem.* **60**, 443–475 (1991).
- Stevens, R. L. & Austen, K. F. Recent advances in the cellular and molecular biology of mast cells. *Immunol. Today* **10**, 381–386 (1989).
- Lindahl, U., Kusche-Gullberg, M. & Kjellén, L. Regulated diversity of heparan sulfate. *J. Biol. Chem.* **273**, 24979–24982 (1998).
- Petersson, I. *et al.* Biosynthesis of heparin: Purification of a 110 kDa mouse mastocytoma protein required for both glucosaminyl N-deacetylation and N-sulfation. *J. Biol. Chem.* **266**, 8044–8049 (1991).
- Orellana, A., Hirschberg, C. B., Wei, Z., Swiedler, S. J. & Ishihara, M. Molecular cloning and expression of a glucosaminoglycan N-acetylglucosaminyl N-deacetylase/N-sulfotransferase from a heparin-producing cell line. *J. Biol. Chem.* **269**, 2270–2276 (1994).
- Eriksson, I., Sandbäck, D., Ek, B., Lindahl, U. & Kjellén, L. cDNA cloning and sequencing of mouse mastocytoma glucosaminyl N-deacetylase/N-sulfotransferase. *J. Biol. Chem.* **269**, 10438–10443 (1994).
- Kusche-Gullberg, M., Eriksson, I., Sandbäck-Pikas, D. & Kjellén, L. Identification and expression in mouse of two heparan sulfate glucosaminyl N-deacetylase/N-sulfotransferase genes. *J. Biol. Chem.* **273**, 11902–11907 (1998).
- Toma, L., Berninson, P. & Hirschberg, C. B. The putative heparin-specific N-acetylglucosaminyl N-deacetylase/N-sulfotransferase also occurs in non-heparin-producing cells. *J. Biol. Chem.* **273**, 22458–22465 (1998).
- Hashimoto, Y., Orellana, A., Gil, G. & Hirschberg, C. B. Molecular cloning and expression of rat liver N-heparan sulfate sulfotransferase. *J. Biol. Chem.* **267**, 15744–15750 (1992).
- Aikawa, J.-i. & Esko, J. D. Molecular cloning and expression of a third member of the heparan sulfate/heparin GlcNAc N-deacetylase/N-sulfotransferase family. *J. Biol. Chem.* **274**, 2690–2695 (1999).
- Lindahl, U., Lidholt, K., Spillman, D. & Kjellén, L. More to “heparin” than anticoagulation. *Thromb. Res.* **75**, 1–32 (1994).
- Salmivirta, M., Lidholt, K. & Lindahl, U. Heparan sulfate: A piece of information. *FASEB J.* **10**, 1270–1279 (1996).
- Rosenberg, R. D., Shworak, N. W., Liu, J., Schwartz, J. J. & Zhang, L. Heparan sulfate proteoglycans of the cardiovascular system. Specific structures emerge but how is synthesis regulated? *J. Clin. Invest.* **99**, 2062–2070 (1997).
- Bullock, S. L., Fletcher, J. M., Beddington, R. S. P. & Wilson, V. A. Renal agenesis in mice homozygous for a gene trap mutation in the gene encoding heparan sulfate 2-sulfotransferase. *Genes Dev.* **12**, 1894–1905 (1998).
- Binari, R. C. *et al.* Genetic evidence that heparin-like glycosaminoglycans are involved in *wingless* signaling. *Development* **124**, 2623–2632 (1997).
- Haerry, T. E., Heslip, T. R., Marsch, J. L. & O'Connor, M. O. Defects in glucuronate biosynthesis disrupt *Wingless* signalling in *Drosophila*. *Development* **124**, 3055–3064 (1997).
- Häcker, U., Lin, X. & Perrimon, N. The *Drosophila sugarless* gene modulates *Wingless* signaling and encodes an enzyme involved in polysaccharide biosynthesis. *Development* **124**, 3565–3573 (1997).
- Huang, C., Sali, A. & Stevens, R. L. Regulation and function of mast cell proteases in inflammation. *J. Clin. Immunol.* **18**, 169–183 (1998).
- Lützelshwab, C., Pejler, G., Aveskog, M. & Hellman, L. Secretory granule proteases in rat mast cells. Cloning of 10 different serine proteases and carboxypeptidase A from various rat mast cell populations. *J. Exp. Med.* **185**, 13–29 (1997).
- Kusche, M., Lindahl, U., Enerbäck, L. & Rodén, L. Identification of oversulphated and galactosaminoglycans in intestinal-mucosal mast cells of rats infected with the nematode worm *Nippostrongylus brasiliensis*. *Biochem. J.* **253**, 885–893 (1988).
- Galli, S. J. & Wershil, B. K. The two faces of the mast cell. *Nature* **381**, 21–22 (1996).
- Fässler, R. & Meyer, M. Consequences of lack of β -1 integrin gene expression in mice. *Genes Dev.* **9**, 1896–1908 (1995).

- Nagy, A., Rossant, J., Nagy, R., Abramow-Newerly, W. & Roder, J. C. Derivation of completely cell culture-derived mice from early-passage embryonic stem cells. *Proc. Natl. Acad. Sci. USA* **90**, 8424–8428 (1993).
- Cheung, W.-F., Eriksson, I., Kusche-Gullberg, M., Lindahl, U. & Kjellén, L. Expression of mouse mastocytoma glucosaminyl N-deacetylase/N-sulfotransferase in human kidney 293 cells results in increased N-sulfation of heparan sulfate. *Biochemistry* **35**, 5250–5256 (1996).
- Lützelshwab, C., Huang, M. R., Kullberg, M. C., Aveskog, M. & Hellman, L. Characterization of mouse mast cell protease-8, the first member of a novel subfamily of mouse mast cell serine proteases, distinct from both the classical chymases and tryptases. *Eur. J. Immunol.* **28**, 1022–1033 (1998).

Acknowledgements. We thank A. Nagy for R1 ES cells, and P. Ekblom and U. Lindahl for helpful suggestions and support. This work was supported by grants from the Swedish Natural Science Research Council, the Swedish Medical Research Council, the European commission, Polysackaridforskning AB, Gustaf V:s 80-årsfond, Magnus Bergvalls stiftelse, Wibergs stiftelse and Stiftelsen Lars Hiertas minne. E.F. is a member of Uppsala University Transgenic Facility, supported by the Swedish Foundation for Strategic Research.

Correspondence and requests for materials should be addressed to L.K. (e-mail: Lena.Kjellen@vmk.slu.se).

The chemokine receptor CCR4 in vascular recognition by cutaneous but not intestinal memory T cells

J. J. Campbell*, G. Haraldsen*†, J. Pan*, J. Rottman‡, S. Qin‡, P. Ponath‡, D. P. Andrew‡, R. Warnke§, N. Ruffing‡, N. Kassam‡, L. Wu‡ & E. C. Butcher*

* Laboratory of Immunology and Vascular Biology, Department of Pathology, Stanford University School of Medicine, Stanford, California 94305, and the Center for Molecular Biology and Medicine, Veterans Affairs Palo Alto Health Care System, Palo Alto, California 94304, USA

‡ LeukoSite Inc., 215 First Street, Cambridge, Massachusetts 02142, USA

§ Department of Pathology, Stanford University School of Medicine, Stanford, California 94305, USA

Lymphocytes that are responsible for regional (tissue-specific) immunity home from the blood to the intestines, inflamed skin or other sites through a multistep process involving recognition of vascular endothelial cells and extravasation¹. Chemoattractant cytokine molecules known as chemokines² regulate this lymphocyte traffic, in part by triggering arrest (stopping) of lymphocytes rolling on endothelium^{3–5}. Here we show that many systemic memory T cells in blood carry the chemokine receptor CCR4 (ref. 6) and therefore respond to its ligands, the chemokines TARC and MDC. These cells include essentially all skin-homing cells expressing the cutaneous lymphocyte antigen and a subset of other systemic memory lymphocytes; however, intestinal ($\alpha\beta 7^+$) memory and naive T cells respond poorly. Immunohistochemistry reveals anti-TARC reactivity of venules and infiltration of many CCR4⁺ lymphocytes in chronically inflamed skin, but not in the gastrointestinal lamina propria. Moreover, TARC induces integrin-dependent adhesion of skin (but not intestinal) memory T cells to the cell-adhesion molecule ICAM-1, and causes their rapid arrest under physiological flow. Our results suggest that CCR4 and TARC are important in the recognition of skin vasculature by circulating T cells and in directing lymphocytes that are involved in systemic as opposed to intestinal immunity to their target tissues.

To identify the circulating lymphocyte subsets that are responsive to different chemokines, we analysed the phenotype of peripheral blood lymphocytes (PBL) migrating to chemokines in a standard trans-well chemotaxis assay^{7,8}. We used flow cytometry to identify natural killer cells (CD56⁺, CD16⁺, CD3⁺), B cells (CD19⁺), and

† Present address: LIIPAT, Institute of Pathology, University of Oslo and Rikshospitalet, Oslo, Norway N-0027

CD8 (CD8⁺, CD27⁺) and CD4⁺ T cells of naive (CD45RA⁺) or memory (CD45RA⁻) phenotype. To determine the role of chemokines in regional lymphocyte trafficking, we also analysed memory-T-cell subpopulations defined by expression of the integrin $\alpha 4\beta 7$, the lymphocyte receptor for the mucosal addressin cell-adhesion molecule (MAdCAM-1); $\alpha 4\beta 7^{\text{hi}}$ memory lymphocytes bind to

vascular MAdCAM-1, move to intestinal sites, and carry memory for intestinal antigens^{1,9,10}, whereas $\alpha 4\beta 7^{\text{lo}}$ cells embody memory for systemic (non-intestinal) immune responses^{1,11}. All lymphocyte subsets apart from natural killer cells responded well to MIP3 β (also called ELC; a ligand for CCR7), and all subsets migrated to SDF-1 (a ligand for CXCR4 (Fig. 1a)). In contrast $\alpha 4\beta 7^{\text{lo}}$ memory

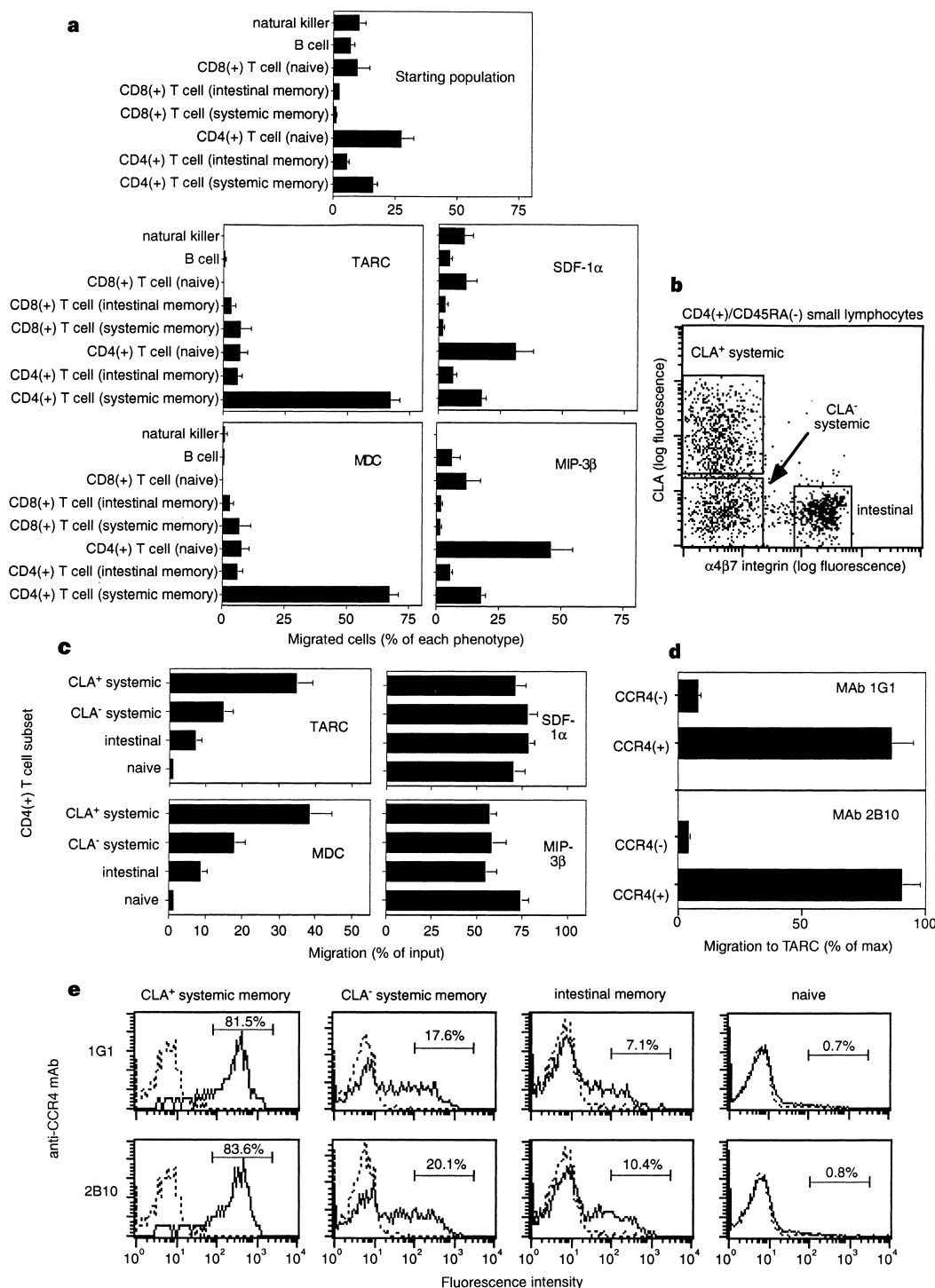


Figure 1 A major subset of systemic memory CD4 T cells (including skin-homing cells) is preferentially attracted by TARC and MDC, and expresses CCR4. **a**, Representation of indicated subsets in unmanipulated compared with specific chemokine-attracted PBL. Mean and s.d.; $n = 3$ experiments. **b**, Gates defining subsets in 1C. **c**, Specific migration of CD4⁺ T subsets. Mean and s.e.m., $n = 8$.

$P < 0.01$ for CLA^{pos} versus gut memory responses to CCR4 ligands. **d**, Comparison of CCR4⁺ versus CCR4⁻ CD4 T-cell migration to TARC, based on staining for monoclonal antibody (mAb) after migration. Representative experiment of two. **e**, FACS analysis of CCR4 expression on CD4⁺ T cells. Dotted line, IgG1 control. One experiment is shown, which is representative of >10 .

CD4⁺ T cells were the predominant population recruited by TARC and by MDC; $\alpha 4\beta 7^+$ memory phenotype CD8⁺ cells were also enriched among lymphocytes responding to these chemokines.

The selective activity of the CCR4 ligands on circulating $\alpha 4\beta 7^+$ memory cells indicated that they might have a role in tissue-selective recruitment of lymphocytes from the blood. We therefore compared the responses of the two most well characterized and antigenically defined tissue-targeted memory-T-cell subsets—intestinal ($\alpha 4\beta 7^{\text{hi}}$) memory CD4⁺ cells and skin-homing memory CD4⁺ T cells (which represent a significant fraction of the systemic $\alpha 4\beta 7^+$ memory population¹). Skin-associated memory T cells are defined by the cutaneous lymphocyte antigen CLA (Fig. 1b), a sialyl Lewis^x-related carbohydrate-dependent epitope that functions as a T-cell homing receptor for inflamed skin^{11–13}. TARC and MDC strongly attract these skin-homing memory cells, whereas intestinal memory T cells respond poorly (Fig. 1c). $\alpha 4\beta 7^+$ CLA⁺ T cells also migrate significantly (although consistently less than the CLA⁺ population), suggesting that the responsiveness of cutaneous memory lymphocytes to CCR4 ligands may be shared with a subset of other systemic memory cells as well. When the responses of CCR4⁺ and CCR4[−] CD4 T cells were compared by staining cells that had migrated with anti-CCR4 monoclonal antibodies, only CCR4⁺ T-helper cells migrated significantly to TARC, indicating that CCR4 expression correlates with TARC-induced migration (D.A. *et al.*, manuscript, in preparation; and Fig. 1d). Consistent with these chemotactic responses, immunofluorescence staining confirmed that CCR4 was differentially expressed on skin-associated and intestinal memory cells: as shown in Fig. 1e, most CLA⁺ memory CD4 cells express high levels of CCR4 (>95% positive; >70% with >100 mean fluorescence units), and $\alpha 4\beta 7^{\text{hi}}$ /CLA[−] cells show intermediate expression (~20% with >100 mean fluorescence units) whereas most $\alpha 4\beta 7^{\text{hi}}$ memory cells are weak or negative (<30% positive, with >90% under 100 mean fluorescence units) B cells (not shown) and naive CD4 cells were negative (D.A. *et al.*, manuscript in preparation). Results were similar when two independent anti-CCR4 antibodies were used. We conclude that CCR4 and its ligands mediate the preferential recruitment of $\alpha 4\beta 7^+$ systemic memory T cells, particularly the skin-homing population.

It has been proposed that chemokines control lymphocyte trafficking not only through stimulation of chemotaxis, but also by triggering rapid integrin-dependent adhesion and arrest of lymphocytes on endothelium³. Consistent with this idea, some chemokines induce rapid arrest of lymphocytes under physiological shear conditions⁴ and can be expressed and/or presented by endothelial cells at sites of lymphocyte extravasation^{5,14–16}. We have found that TARC is expressed by activated endothelium: TARC (but not MDC) messenger RNA is readily detected in northern-blot analysis of RNA from endotoxin- and cytokine-stimulated human umbilical vein endothelial cells (J.P., unpublished results) and in primary cultures of enriched human-tonsil HEV cells (isolated as described¹⁷; results not shown). As TARC is expressed by endothelial cells, it could be available to regulate vascular interactions with lymphocytes.

To test whether TARC is presented by venules at sites of CCR4⁺ T-cell homing, we immunohistologically assayed the presentation of TARC by venules, and of CCR4 by infiltrating T cells, in biopsies of chronically inflamed skin from patients with a variety of dermatological disorders. Anti-TARC monoclonal antibody (mAb) reacted with venules associated with lymphocyte recruitment in inflamed skin (Fig. 2a–c, g), including (but not limited to) most E-selectin-expressing venules. CCR4-expressing cells were abundant in these lesions (Fig. 2d insert, for example). Many venules in minimally inflamed and uninfamed areas of the dermis showed significant but lesser anti-TARC reactivity as well. Reactivity was also found on many high endothelial venules in inflamed tonsils, which mediate the recruitment of subsets of memory as well as naive lymphocytes^{1,10} and so may express several chemokines. Venules involved in lymphocyte recruitment to gastrointestinal lamina propria (identified by MAdCAM-1 staining in biopsies of small intestine, stomach, and normal and inflamed colon) were predominantly negative (Fig. 2e–g). Lamina propria lymphocytes were mostly CCR4[−] (Fig. 2f, insert). Parallel studies of macaque skin revealed scattered constitutive anti-TARC reactivity of venules, but with more extensive reactivity in experimentally induced delayed-type-hypersensitivity reactions¹⁸. The results indicate that TARC is well positioned on inflamed endothelium in skin to help control the

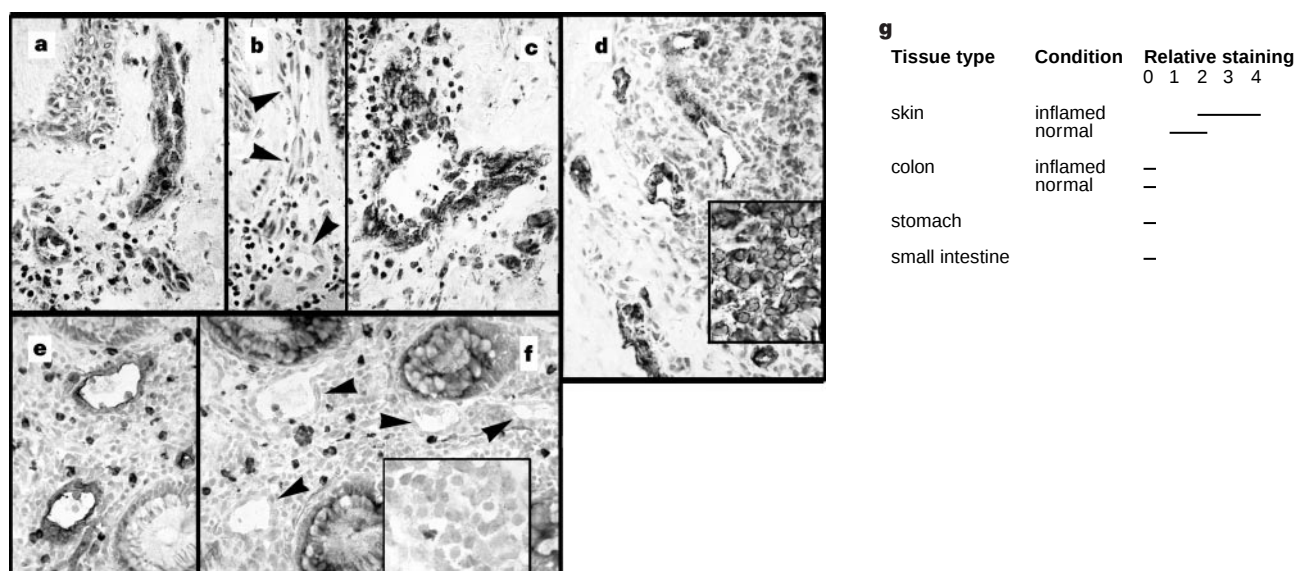


Figure 2 Anti-TARC reactive venules and infiltrating CCR4⁺ cells in inflammatory skin lesions. Juxta-epidermal vessels in psoriasis stained with anti-TARC (a) but not by control IgG1 mAb (arrows) (b). Positive dermal venules in psoriasis (c) and lichen planus (d), with higher power magnification showing an intensely CCR4⁺

lymphocytic infiltrate (d, inset). In contrast, MAdCAM-1-positive (e) venules in the lamina propria in an inflamed colon (Crohn's disease) do not react with anti-TARC mAb (arrows) (f); and infiltrating lymphocytes are CCR4[−] (f, inset). g, Summary of endothelial anti-TARC reactivity in various tissues (see Methods).

vascular arrest and adhesion of circulating CCR4⁺ lymphocytes. Other systemic tissues remain to be examined, but the inflamed brains of patients with multiple sclerosis have failed to reveal anti-TARC-reactive endothelium or CCR4 expression by infiltrating T cells¹⁹.

To test whether TARC could trigger rapid integrin activation and integrin-dependent arrest of circulating lymphocytes, we assayed for rapid adhesion to ICAM-1 and found that CD4⁺ populations enriched in CLA⁺ (skin) or $\alpha 4\beta 7^-$ (intestinal) memory cells responded equally well to SDF-1, but that only the CLA-enriched subset was effectively triggered by TARC (Fig. 3a). The inability of intestinal memory cells to respond to TARC by rapid adhesion, which contrasts with their low but detectable chemotaxis to TARC, may reflect the low levels of CCR4 on the positive subset of intestinal memory cells (median fluorescence was ~7-fold lower than for CLA⁺ T cells): thus CCR4 expression may be insufficient to generate a robust pro-adhesive response, which typically requires the engagement of many receptors⁷.

Triggered adhesion of lymphocytes to endothelium *in vivo* must occur within seconds and under conditions of strong shear stress at the vascular wall. In inflamed skin, the vascular CLA receptor E-selectin^{13,20} mediates initial cell contact (tethering) and supports rolling; $\beta 2$ -integrin ligands mediate arrest (with variable contributions from $\alpha 4\beta 1$ integrin and VCAM-1)^{1,18,20}. To determine whether TARC can mediate activation-dependent arrest of lymphocytes rolling under physiological conditions, we coated capillary tubes with E-selectin in combination with TARC and/or ICAM-1, and passed PBL through the tube at a wall shear stress of 2 dynes cm⁻²

(under these conditions, the mean velocity of rolling cells was 5.1 $\mu\text{m s}^{-1}$ (± 1.9 s.d.) on E-selectin plus ICAM-1, without chemokine). We found that many rolling lymphocytes came to a rapid stop (Fig. 3b). Arrest occurred within less than 1 second (often within 30 ms) after initial rolling for most stopping cells (Fig. 3c) and required the combined presence of all three components.

A model emerges in which CLA and CCR4, participating in sequential events in the specialized adhesion/activation cascade in inflamed skin, can each make an essential contribution to the specificity of lymphocyte arrest and recruitment. Although CLA expression (conferring tethering and rolling on vascular E-selectin) is unique among T cells, other leukocytes (for example neutrophils and monocytes) also roll on E-selectin. Monocytes and granulocytes, however, do not respond to TARC²¹, so their interactions with E-selectin-positive/TARC-positive venules in chronically inflamed skin should be unproductive and only rolling T cells should be made to stop. A requirement for CCR4 could also enhance the selectivity of cutaneous memory-T-cell recruitment in settings in which vascular tethering molecules for other T-cell subsets, such as P-selectin or L-selectin binding sites, are induced in conjunction with the CLA ligand E-selectin. Thus, CLA and CCR4 may act as independent discriminators that work together (in conjunction with $\alpha 4\beta 1$ and LFA-1) to provide the high specificity observed *in vivo*, illustrating how combinatorial control is possible in multistep leukocyte trafficking³. Other chemokines like MDC and their receptors may be involved in cutaneous homing events too, especially in diapedesis, and CCR4 may participate in homing/adhesion cascades in a subset of other systemic inflammatory sites.

CCR4 is expressed by *in vitro*-polarized Th2 T lymphocytes²²⁻²⁴; however, the same *in vitro* polarization protocols induce selective expression of CLA on Th1 but not Th2 cells²⁵, contrasting with the coordinated expression of CLA and CCR4 on circulating cutaneous memory cells in blood. CCR4 can also be expressed by non-Th2 T-cell lines generated in the presence of TGF- β ²⁶, or even by some Th1 cells *in vitro*²⁷. Our results are consistent with this independent regulation of CCR4 and the cytokine profile, in that CLA⁺ T cells (shown here to be CCR4⁺) comprise cells with both Th1 and Th2 potential²⁸. Moreover, >35% of circulating CD4⁺/CLA⁺/CCR4⁺ T cells co-express the proposed *in vitro* Th1-associated^{26,29} chemokine receptor CXCR3, and >50% stain with the putative Th1-associated^{26,29} CCR5 (D.P.A. and L.W., manuscript in preparation). We propose that CCR4 on circulating memory lymphocytes can function as a chemoattractant 'homing receptor' for endothelium, independently of cytokine commitments. Its expression on differentiated effector cells *in vitro* may imply a parallel role for the relatively infrequent effector cells that enter the blood, or may relate to the role(s) of CCR4 in the migration and behaviour of activated T cells following extravasation at sites of inflammation.

We conclude that TARC and its lymphocyte receptor CCR4 are involved in the homing of a major subset of circulating systemic memory T cells, and that they participate in memory-T-cell interactions with vascular endothelium at cutaneous sites of inflammation. The involvement of CCR4 and its chemokine ligands in vascular recognition in other systemic sites, and the nature and trafficking properties of the subset of CCR4⁺ but CLA⁻CD4 cells in blood (representing ~9% of the total memory CD4 pool) need to be investigated. Our findings provide evidence for the involvement of chemokines in tissue-selective lymphocyte-endothelial recognition during inflammation and indicate that CCR4 and its ligands are important in the regional targeting of memory T cells and thus in the functional segregation of intestinal and systemic (especially cutaneous) immune responses. □

Methods

Chemotaxis assays. Migration of human peripheral blood lymphocytes through 5- μm pores and calculation of per cent migration of cells of the indicated phenotypes, based on FACS analysis of migrated compared with

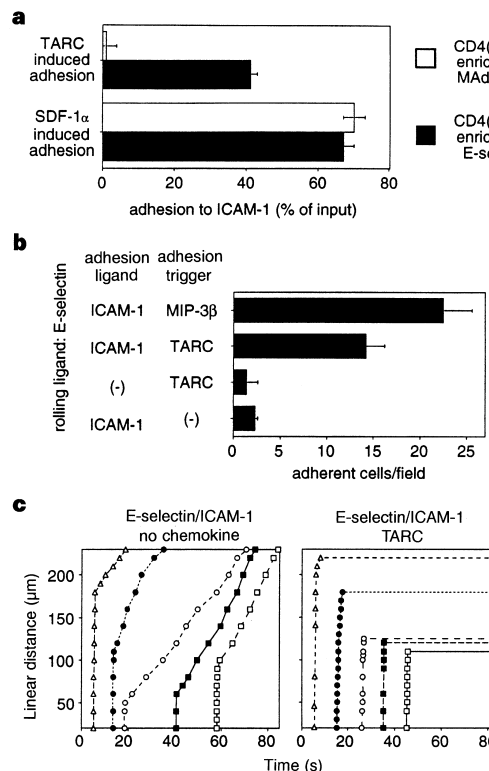


Figure 3 TARC and rapid ICAM-1 adhesion. **a**, TARC triggers adhesion of skin-associated (but not intestinal) memory CD4 T cells in a static adhesion assay; the background adhesion (~10%) is subtracted. Representative of two donors. Error bars: range of duplicates. **b**, Adherent lymphocytes accumulate at ~2.0 dynes cm⁻² shear on a capillary tube coated with E-selectin, chemokine and ICAM-1. Error bars: s.d. for four donors, 10 fields each. **c**, TARC-induced ICAM-1 adhesion of lymphocytes rolling on E-selectin is extremely rapid. Each line depicts the behaviour of an individual cell interacting with the tubes shown in **b**.

input cells (for example, CLA⁺ or CCR4⁺CD4 subsets; Fig. 1c, d), were determined as described⁸. 4 (SDF-1 α and MIP-3 β), 16 (TARC and MDC) or 24 (medium control) replicate wells were used per chemokine per donor. To calculate specific migration, the mean number of migrated cells per well in the absence of chemokine (background) was subtracted from the total number of cells per well that migrated to chemokine in parallel wells. To calculate the percent representation of each lymphocyte subtype in the migrated populations (Fig. 1a), the number of specifically migrated cells per well of each subtype was divided by the total number of specifically migrated cells per well. In Fig. 1d, results are normalized and expressed as a percentage of maximal migration. Chemokines were titred and used at the optimal concentrations for chemotaxis of unfractionated lymphocytes as follows. TARC, MDC and SDF-1 α : 100 nM; MIP-3 β , 1 μ M. Synthetic human TARC and SDF-1 α were gifts from M. Siani and D. Thompson, Gryphon Sciences. Recombinant human MDC was a gift from Amgen (D.P.A.) and MIP-3 β was from PeproTech EC.

Directly conjugated antibodies. For use in FACS analysis, directly conjugated antibodies were obtained from PharMingen unless indicated otherwise. These were: fluorescein isothiocyanate (FITC)-conjugated anti-CD3 (clone UCHT1), CD45RA (clone HI100), CLA (clone HECA 452; E.C.B.); phycoerythrin (PE)-conjugated anti-CD56 (clone B159), CD19 (clone B43), CD27 (M-T271), α 4 β 7 (clone ACT-1; LeukoSite); biotinylated anti-CD45RA (clone HI100), CD16 (clone 3G8), CD8 (clone RPA-T8); APC-conjugated anti-CD4 (clone RPA-T4), CD8 (clone RPA-T8), CD56 (clone B159), TCR α β (clone T10B9.1A-31). Unconjugated CCR4 mAbs were detected with a biotinylated horse-anti-mouse IgG (Vector Labs). The second-stage reagent used for all biotinylated antibodies was streptavidin-PerCP (Beckton Dickinson).

Unconjugated mAbs. Anti-CCR4 mAbs 1G1 and 2B10 (D.P.A. and L.W., manuscript in preparation) were independently produced against CCR4-transfected mouse L1/2 cells as described³⁰. Both stained with CCR4-expressing L1/2 cells, but not with controls expressing other chemokine and related receptors (including CXCR1–5, CX3CR1, CCR1–3, CCR5–9, Bob and Bonzo). Anti-TARC mAb LS142-2D8 was produced against synthetic human TARC by standard techniques, and selected by ELISA for reactivity with TARC but not MDC or other chemokines (S.Q., manuscript in preparation). In addition to the endothelial reactivity described here, LS142-2D8 crossreacts with smooth muscle cells in tissue sections (not shown).

Immunohistology. Frozen sections were fixed for 10 min at room temperature in 4% paraformaldehyde in phosphate-buffered saline (PBS). After washing in PBS, sections were stained with standard immunoperoxidase²⁰ followed by haematoxylin. Anti-TARC (LS142-2D8) staining was compared with IgG1 control on each section, and graded for reactivity on the basis of the average venular intensity. The most intense staining was in cutaneous biopsies of psoriasis and lichen planus (3 to 4⁺), with significant staining in 2 cases of mild atopic dermatitis (2⁺), in nonspecific dermatitis (2⁺), in uninvolved areas of several of the above biopsies (generally 1⁺, with some 0 and 2⁺ venules); in uninvolved areas of several of the above biopsies (generally 1⁺, with some 0 and 2⁺ venules); and in two primate DTH samples (1–2⁺) and control primate skin (1⁺). In contrast, lamina propria venules in stomach ($n = 2$), normal colon ($n = 2$), colon in Cohn's disease ($n = 2$) or ulcerative colitis ($n = 2$) or small intestine ($n = 3$) were predominantly negative, or gave only infrequent focal reactivity (0).

Static rapid adhesion assay of skin- and gut-associated memory CD4⁺ T cells. CD4⁺ cells were purified from fresh PBL as described^{4,8}, then incubated on fibroblasts transfected with MAdCAM-1 or E-selectin in T-175s at 37 °C for 30 min. Non-bound cells were washed away with warm complete medium, bound cells were recovered by elution with HBSS buffer free of divalent cations for E-selectin or with 0.5 mM EDTA for MAdCAM-1. About 70% of the E-selectin-adherent population was CLA⁺, and only 6% were α 4 β 7^{hi} (with the remainder being naive and α 4 β 7^{lo} memory cells); the MAdCAM-1-adherent population consisted of ~60% α 4 β 7^{hi} memory T cells and only ~5% CLA⁺ cells. Static rapid adhesion to ICAM-1 was assayed as described^{4,7}. Adherent cells were counted 2 min after exposure to chemokine.

Rapid ICAM-1 adhesion of lymphocytes rolling on E-selectin under shear. Fresh human peripheral blood lymphocytes (99.8% pure and depleted of monocytes) were prepared and assayed for interaction and arrest under shear in capillary tubes coated with the indicated adhesion molecules and chemokines, as described⁴.

Purified adhesion molecules. Human E-selectin was purified from E-selectin-transfected L1/2 cells on an affinity column of monoclonal antibody E8.16-3 (gift from E. Berg) conjugated to Sepharose⁴. ICAM-1 was prepared as described⁴.

Received 28 April; accepted 25 June 1999.

- Butcher, E. C., Williams, M., Youngman, K., Rott, L. & Briskin, M. Lymphocyte trafficking and regional immunity. *Adv. Immunol.* **72**, 209–253 (1999).
- Zlotnik, A., Morales, J. & Hedrick, J. A. Recent advances in chemokines and chemokine receptor. *Crit. Rev. Immunol.* **19**, 1–47 (1999).
- Butcher, E. C. Leukocyte-endothelial cell recognition: three (or more) steps to specificity and diversity. *Cell* **67**, 1033–1036 (1991).
- Campbell, J. J. et al. Chemokines and the arrest of lymphocytes rolling under flow conditions. *Science* **279**, 381–384 (1998).
- Gunn, M. D. et al. A chemokine expressed in lymphoid high endothelial venules promotes the adhesion and chemotaxis of naive T lymphocytes. *Proc. Natl Acad. Sci. USA* **95**, 258–263 (1998).
- Power, C. A. et al. Molecular cloning and functional expression of a novel CC chemokine receptor cDNA from a human basophilic cell line. *J. Biol. Chem.* **270**, 19495–19500 (1995).
- Campbell, J. J., Qin, S., Bacon, K. B., Mackay, C. R. & Butcher, E. C. Biology of chemokine and classical chemoattractant receptors: differential requirements for adhesion-triggering versus chemotactic responses in lymphoid cells. *J. Cell Biol.* **134**, 255–266 (1996).
- Campbell, J. J. et al. 6-C-kine (SLC), a lymphocyte adhesion-triggering chemokine expressed by high endothelium, is an agonist for the MIP-3 β receptor CCR7. *J. Cell Biol.* **141**, 1053–1059 (1998).
- Mackay, C. R., Andrew, D. P., Briskin, M., Ringle, D. J. & Butcher, E. C. Phenotype and migration properties of three major subsets of tissue homing T cells in sheep. *Eur. J. Immunol.* **26**, 2433–2439 (1996).
- Williams, M. B. & Butcher, E. C. Homing of naive and memory T lymphocyte subsets to Peyer's patches, lymph nodes, and spleen. *J. Immunol.* **159**, 1746–1752 (1997).
- Santamaria-Babi, L. F. et al. Circulating allergen-reactive T cells from patients with atopic dermatitis and allergic contact dermatitis express the skin-selective homing receptor, the cutaneous lymphocyte-associated antigen. *J. Exp. Med.* **181**, 1935–1940 (1995).
- Pickar, L. J., Michie, S. A., Rott, L. S. & Butcher, E. C. A unique phenotype of skin-associated lymphocytes in humans. Preferential expression of the HECA-452 epitope by benign and malignant T cells at cutaneous sites. *Am. J. Pathol.* **136**, 1053–1068 (1990).
- Berg, E. L. et al. The cutaneous lymphocyte antigen is a skin lymphocyte homing receptor for the vascular lectin endothelial cell-leukocyte adhesion molecule 1. *J. Exp. Med.* **174**, 1461–1466 (1991).
- Anderson, A. O. & Shaw, S. T-cell adhesion to endothelium: the FRC conduit system and other anatomic and molecular features which facilitate the adhesion cascade in lymph node. *Semin. Immunol.* **5**, 271–282 (1993).
- Middleton, J. et al. Transcytosis and surface presentation of IL-8 by venular endothelial cells. *Cell* **91**, 385–395 (1997).
- Utgard, J. O., Jahnsen, F. L., Bakka, A., Brandtzaeg, P. & Haraldsen, G. Rapid secretion of prestored interleukin 8 from Weibel-Palade bodies of microvascular endothelial cells. *J. Exp. Med.* **188**, 1751–1756 (1998).
- Baekkevold, E. S. et al. Culture characterization of differentiated high endothelial venule cells from human tonsils. *Lab. Invest.* **79**, 327–336 (1999).
- Silber, A. et al. Recruitment of lymphocytes during cutaneous delayed hypersensitivity in nonhuman primates is dependent on E-selectin and vascular cell adhesion molecule 1. *J. Clin. Invest.* **93**, 1554–1563 (1994).
- Balashov, K. E., Rottman, J. B., Weiner, H. L. & Hancock, W. W. CCR5(+) and CXCR3(+) T cells are increased in multiple sclerosis and their ligands MIP-1 α and IP-10 are expressed in demyelinating brain lesions. *Proc. Natl Acad. Sci. USA* **96**, 6873–6878 (1999).
- Pickar, L. J., Kishimoto, T. K., Smith, C. W., Warnock, R. A. & Butcher, E. C. ELAM-1 is an adhesion molecule for skin-homing T cells. *Nature* **349**, 796–799 (1991).
- Imai, T., Baba, M., Nishimura, M., Takagi, S. & Yoshie, O. The T cell-directed CC chemokine TARC is a highly specific biological ligand for CC chemokine receptor 4. *J. Biol. Chem.* **272**, 15036–15042 (1997).
- Bonocchi, R. et al. Differential expression of chemokine receptors and chemotactic responsiveness of type 1 T helper cells (Th1s) and Th2s. *J. Exp. Med.* **187**, 129–134 (1998).
- Andrew, D. P. et al. STCP-1 (MDC) CC chemokine acts specifically on chronically activated Th2 lymphocytes and is produced by monocytes on stimulation with Th2 cytokines IL-4 and IL-13. *J. Immunol.* **161**, 5027–5038 (1998).
- Sallusto, F., Lenig, D., Mackay, C. R. & Lanzavecchia, A. Flexible programs of chemokine receptor expression on human polarized T helper 1 and 2 lymphocytes. *J. Exp. Med.* **187**, 875–883 (1998).
- Austrup, F. et al. P- and E-selectin mediate recruitment of T-helper-1 but not T-helper-2 cells into inflamed tissues. *Nature* **385**, 81–83 (1997).
- Sallusto, F., Lanzavecchia, A. & Mackay, C. R. Chemokines and chemokine receptors in T-cell priming and Th1/Th2-mediated responses. *Immunol. Today* **19**, 568–574 (1998).
- D'Ambrosio, D. et al. Selective up-regulation of chemokine receptors CCR4 and CCR8 upon activation of polarized human type 2 Th cells. *J. Immunol.* **161**, 5111–5115 (1998).
- Teraki, Y. & Pickar, L. J. Independent regulation of cutaneous lymphocyte-associated antigen expression and cytokine synthesis phenotype during human CD4⁺ memory T cell differentiation. *J. Immunol.* **159**, 6018–6029 (1997).
- Qin, S. et al. The chemokine receptors CXCR3 and CCR5 mark subsets of T cells associated with certain inflammatory reactions. *J. Clin. Invest.* **101**, 746–754 (1998).
- Qin, S. et al. Expression of monocyte chemoattractant protein-1 and interleukin-8 receptors on subsets of T cells: correlation with transendothelial chemotactic potential. *Eur. J. Immunol.* **26**, 640–647 (1996).

Acknowledgements. We thank L. Rott for important advice on flow cytometry and cell separation, E. Resurreccion for immunohistology, and E. Pfendt for frozen tissue archiving. Some tissues were obtained from the National Disease Research Interchange, a service organization funded by the NIH, or from the Human Cooperative Tissue Network. This work was supported by grants from the NIH and an Award from the Department of Veterans Affairs (E.C.B.) and by the FACS Core Facility of the Stanford Digestive Disease Center. G.H. was supported by the Norwegian Cancer Society, J.J.C. was supported by a grant from the NIH Cancer Etiology, Prevention, Detection, and Diagnosis, and by an NIH Individual National Research Service award. J.J.C. and J.P. are recipients of postdoctoral fellowships from the Arthritis Foundation.

Correspondence should be addressed to J.J.C. (e-mail: jcampbel@cmgm.stanford.edu) and requests for materials to D.P.A. (e-mail: david_andrew@leukosite.com) or L.W. (e-mail: lijun_wu@leukosite.com).

Binding of phytochrome B to its nuclear signalling partner PIF3 is reversibly induced by light

Min Ni, James M. Tepperman & Peter H. Quail

Department of Plant and Microbial Biology, University of California, Berkeley, California 94720, and USDA/ARS Plant Gene Expression Center, 800 Buchanan Street, Albany, California 94710, USA

The phytochrome photoreceptor family directs plant gene expression by switching between biologically inactive and active conformers in response to the sequential absorption of red and far-red photons^{1,2}. Several intermediates that act late in the phytochrome signalling pathway have been identified, but fewer have been identified that act early in the pathway^{3,4}. We have cloned a nuclear basic helix-loop-helix protein, PIF3, which can bind to non-photoactive carboxy-terminal fragments of phytochromes A and B and functions in phytochrome signalling *in vivo*⁵. Here we show that full-length photoactive phytochrome B binds PIF3 *in vitro* only upon light-induced conversion to its active form, and that photoconversion back to its inactive form causes dissociation from PIF3. We conclude that photosensory signalling by phytochrome B involves light-induced, conformer-specific recognition of the putative transcriptional regulator PIF3, providing a potential mechanism for direct photoregulation of gene expression.

The phytochromes are chromoproteins encoded by a gene family of five members, *PHYA* to *PHYE*, in *Arabidopsis*⁶. Each photoreceptor molecule consists of a dimer of two ~125K subunits with a single covalently linked tetrapyrrole chromophore. The 125K polypeptide folds into two major structural domains, each encompassing about half of the molecule: an amino-terminal domain that cradles the chromophore, and a carboxy-terminal domain that mediates dimerization⁷. Changes in gene expression are triggered by photoconversion of the photoreceptor to its active (Pfr) form^{8,9}.

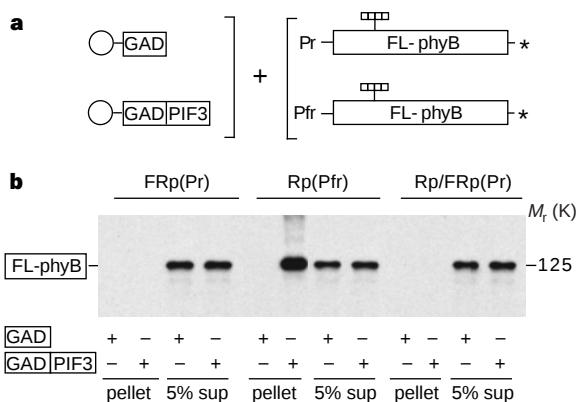


Figure 1 Phytochrome B (phyB) binds to PIF3 specifically as the biologically active conformer Pfr. **a**, Experimental design. Recombinant Gal4-activation domain (GAD) or GAD-PIF3 fusion protein, immobilized on beads (represented by circles) were each mixed with radioactively labelled (asterisks), chromophore-ligated (represented as small striped bars), full-length (FL) phyB, which was converted to either the inactive (Pr) or Pfr form by 5-min pulses of far-red or red irradiation, respectively. **b**, Autoradiography, of pelleted phyB and 5% of supernatant (5% sup) obtained by centrifugation after 2 h incubation in the dark at 4°C, in the presence (+) or absence (-) of either GAD or GAD-PIF3, after pulse irradiation with 5 min far-red light (FRp), 5 min red light (Rp), or with 5 min red followed immediately by 5 min of far-red light (Rp/FRp).

Several components have been cloned that are involved in signal transduction beyond the point of convergence of the phytochrome and blue/ultraviolet-light-receptor pathways^{4,10,11}. PIF3, a putative transcription factor and phytochrome signalling partner, was isolated by virtue of its interaction with the non-photoactive C-terminal domains of phytochromes A and B (ref. 5). We now investigate whether this binding reaction is photoregulated in the full-length photoactive phytochrome B molecule.

We synthesized radioactively labelled full-length *Arabidopsis* phytochrome B by *in vitro* transcription-translation, ligated the chromophore to the protein, and measured the binding of the holoprotein to matrix-immobilized PIF3 in response to pulse irradiations. Figure 1 shows that a 5-min pulse of red, but not far-red, light induces binding of phytochrome B to PIF3. Five minutes of far-red irradiation immediately after an initial red pulse reversed the red-light effect, so that no binding of phytochrome B to PIF3 was detectable. These results indicate that the binding of full-length phytochrome B to PIF3 is photoregulated, occurring predominantly, if not exclusively, in the biologically active Pfr form.

As this experiment involved a two-hour incubation in the dark after pulse irradiation, it was not clear whether giving a far-red pulse after a red pulse prevented phytochrome B from binding or caused it to dissociate after binding. We therefore followed the time course of red-pulse-induced binding and the response of bound phytochrome B to a delayed far-red pulse. Figure 2 shows that binding was initially rapid during the dark incubation period at 4°C after the red-light pulse, then increased more slowly up to 4 h. A far-red pulse given 2 h after the initial red pulse caused a rapid decrease in bound phytochrome B to background levels (Fig. 2). These results show that phytochrome B can bind to PIF3 in darkness after pulse photoconversion to its active form, without the need for continuous photoactivation, and that once bound, reconversion to its inactive Pr form causes the rapid dissociation of phytochrome from PIF3.

We have shown previously that a set of missense mutations clustered in the C-terminal domains of phytochromes A and B

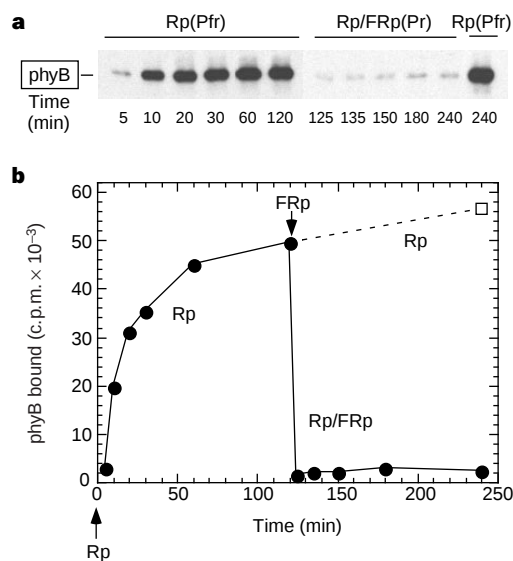


Figure 2 Induced binding of the active (Pfr) form of phytochrome B to PIF3 is reversed by photoconversion back to the inactive Pr form. **a**, Autoradiograph showing the time course of full-length phyB association with matrix-immobilized GAD-PIF3. Samples were irradiated with a 5-min pulse of red light (Rp), incubated in darkness at 4°C for the periods indicated up to 120 minutes, and then either incubated further to 240 min or given a pulse of 5 min of far-red light (Rp/FRp) and incubated for the periods indicated up to 240 min before centrifugation. **b**, Quantitative determination of the amounts of PIF3-bound, radioactively labelled phyB shown in **a**.

interferes with the transfer of signals from the photoreceptor to transduction pathway components *in vivo*, without affecting photo-signal perception^{2,12,13}. Figure 3 shows that the three missense mutations tested here impair binding of photoactivated phytochrome to PIF3. These results are consistent with the conclusion that the reduced signalling by the mutant proteins *in vivo* is caused by a weakened interaction with PIF3, perhaps as a result of a reduction in binding affinity.

We investigated the relative functional importance of the N- and C-terminal domains in the interaction with PIF3 by comparing the binding to PIF3 of full-length phytochrome B with that of each separate domain, as well as with the apoprotein. Figure 4 shows that PIF3 binds strongly to photoactivated full-length phytochrome B but only moderately to the phytochrome B N-terminal domain and comparatively weakly to the C-terminal domain. These results indicate that there are binding determinants in each half of the protein that are strongly synergistic in the full-length molecule. This synergism requires the chromophore and conversion to the active conformer, because it is not seen in the full-length apoprotein or in the inactive form of phytochrome B, indicating that photoconversion to the active form drives the creation of a high-affinity binding site(s) from elements of each domain of the photoreceptor. It is unclear whether these elements are two or more separate sequences in the polypeptide or the residual 'halves' of a disrupted single site that would normally span the junction between the two domains. The red-light-induced binding of the isolated N-terminal domain to PIF3 indicates that the binding element in this domain undergoes a conformational change upon photoactivation. It is not known whether the C-terminal element can also undergo such a photo-induced switch in conformation when coupled to the chromophore-bearing domain in the full-length photoreceptor, or whether it remains constitutive.

The interaction of PIF3 with the C-terminal domain of phytochrome B, which appeared to be substantial in a previous pull-down experiment⁵ seems to be relatively weak in our present assay.

Preliminary evidence indicates that the lower concentrations of phytochrome B used in the binding reactions described here may explain this difference (data not shown).

Our results support the conclusion that PIF3 is a primary signalling partner of phytochrome B. At least three steps in the transduction process from phytochrome B to PIF3 could be reversibly dependent on the light-induced formation of the biologically active conformer of phytochrome B: the translocation of phytochrome B from the cytoplasm to the nucleus; binding of phytochrome B to PIF3; and the biochemical trans-acting constituting signal transfer to PIF3. Phytochrome B is induced to translocate from the cytoplasm to the nucleus when *Arabidopsis* seedlings are exposed to red light¹⁴. This would regulate access to PIF3, which may be constitutively nuclear⁵, and indicates that photoactivated phytochrome B interacts specifically with nuclear import components. Our results show that phytochrome B is rapidly induced to associate with, and subsequently dissociate from, PIF3 upon sequential photoconversion to its active and inactive forms, respectively. The rapid reversible binding of the two partners caused by this photo-induced switching between the active and inactive states of the photoreceptor could be critical for regulating signal flow to PIF3. There is evidence for autophosphorylation of phytochrome A, suggesting that the phytochromes may be photoregulated serine/threonine kinases¹⁵. We do not know whether signal transfer from phytochrome B to PIF3 involves a photoregulated biochemical reaction such as *trans*-phosphorylation, or an allosteric effect on some activity such as DNA binding. Nevertheless, several steps in the primary phytochrome signal transduction process are apparently subject to photoregulation.

Our results verify that the phytochrome molecule functions as a unique binary optical storage device, whose biochemical output is controlled by the stored information. Because PIF3 is a putative transcriptional regulator, it seems likely that this biochemical output directly regulates target gene expression. □

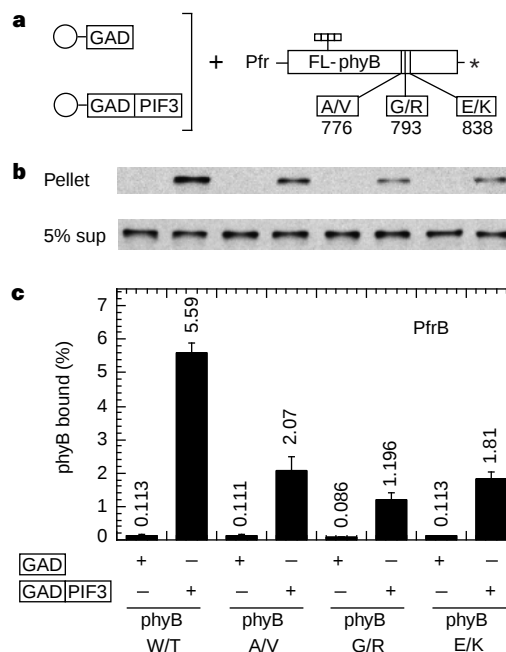


Figure 3 Missense mutations in phytochrome B impair PIF3 binding. **a**, Experimental design. Recombinant Gal4 activation domain (GAD) or GAD-PIF3 fusion protein was immobilized on beads (circles), then mixed with radioactively labelled (asterisk), chromophore-ligated (striped bar) wild-type full-length (FL) phyB in the Pfr form, or its mutant derivatives (A776/V, G793/R or E838/K) and

incubated for 20 min at 4°C in darkness before centrifugation. **b**, Autoradiography of pelleted phyB and 5% of supernatant fractions. **c**, Quantification of PIF3-bound, radioactive phyB shown in **b**, expressed as a per cent of total in supernatant and pellet. Each histogram corresponds to the sample above it in **b**, and represents the mean $\pm$ standard error of three experiments.

Methods

Construction of expression vectors. To produce bait proteins in *Escherichia coli*, the Gal4 activation domain (GAD) sequence was amplified by PCR using primers containing 5' *NdeI* and 3' *Sall* restriction-enzyme cutting sites and cloned into the *NdeI/XhoI* sites of the pRSETB vector (Invitrogen). The expression vector construct for GAD–PIF3 was made as described⁵. To make T7-promoter driven *in vitro* transcription/translation templates, the full-length phytochrome B gene (*PHYB*) coding sequence was amplified by PCR and cloned into the pET-3c vector as described¹⁶. Full-length *PHYB* sequences bearing missense mutations were made by using a QuikChange site-directed mutagenesis kit with wild-type *PHYB* as template (Stratagene). The *PHYB* N-terminal sequence (corresponding amino acids 1–645) was amplified by PCR using primers containing 5' *NdeI* and 3' *EcoRI* restriction-enzyme cutting

sites and cloned into *NdeI/EcoRI* sites of the pRSETC vector (Invitrogen). The *PHYB* C-terminal sequence corresponding to amino acids 645–1,211) was amplified by PCR and cloned into the pET-15b vector as described⁵.

Expression of bait and prey proteins. Both GAD and GAD–PIF3 fusion proteins were expressed in *E. coli* host-strain BL21 cells. Overnight *E. coli* cultures were diluted 20-fold into fresh LB medium containing 50 $\mu\text{g ml}^{-1}$ ampicillin, grown at 30 °C until they reached an absorbance at 600 nm of 0.6, and induced with 0.5 mM isopropyl β -D-thiogalactopyranoside (IPTG) for 3 h at 30 °C. Total protein was isolated by suspending the cells in 3 vol of PBS buffer (pH 7.4) containing 1 mM PMSE, 2 mM EDTA, 0.1% 2-mercaptoethanol, 5 mM benzimidazole and 0.5% NP-40. The cell suspension was then sonicated for 4 $\times$ 20 s on ice with a 40-watt Tekmar sonic disruptor and cleared at 15,000g for 10 min at 4 °C. Total protein was analysed by SDS–PAGE and the expression of GAD or GAD–PIF3 was monitored and quantified by immunoblotting using a monoclonal antibody against the Gal4 activation domain (Santa Cruz Biotechnology). Full-length *PHYB*, and N- and C-terminal domains of *PHYB* were synthesized and labelled with ³⁵S-Met in the T'nT *in vitro* transcription/translation system (Promega). The chromophore was attached to full-length or the N-terminal domain of *PHYB* by adding phycocyanobilin to the T'nT reaction at 10 μM final concentration and incubating in darkness for 1 h at 4 °C. The phycocyanobilin was crude methanolsate, prepared as described¹⁷.

Preparation of bait agarose beads. GAD and GAD–PIF3 were attached to agarose beads by precipitating 20 or 200 μg total protein of *E. coli* lysates containing similar levels of GAD and GAD–PIF3 using 1 μg monoclonal antibody against GAD and 20 μl protein-A agarose beads (Santa Cruz Biotechnology) incubated for 2 h at 4 °C in 0.3 ml PBS binding buffer (pH 7.4) as described⁵. The beads, with attached GAD or GAD–PIF3 proteins, were washed twice with PBS binding buffer without BSA and PMSE, and recovered by centrifugation at 3,000 r.p.m. for 5 min at 4 °C.

In vitro interaction assay. Preparations of the prey molecules containing full-length, N-, or C-terminal domains of *PHYB* synthesized in the T'nT reaction were first precleared by adding 20 μl protein A/G plus agarose beads for 1 h in 0.3 ml PBS binding buffer. Precleared supernatants were then mixed with the bait agarose beads bearing equal amounts of GAD or GAD–PIF3. The immobilized bait proteins and bound phytochrome B proteins were present in roughly equimolar amounts in these assays. The mixtures were set on ice and received 5 min of red or far-red light illumination using an LED light source (QBEAM 2200) with a red light output of 74 $\mu\text{mol s}^{-1} \text{m}^{-2}$ at 664 nm or a far-red light output of 84 $\mu\text{mol s}^{-1} \text{m}^{-2}$ at 748 nm (Quantum Devices). The mixture was then incubated at 4 °C in darkness for different times, as specified in the figure legends. Agarose beads were then pelleted, washed four times with PBS binding buffer without BSA and PMSE, and the attached proteins were solubilized and analysed by 10% SDS–PAGE as described⁵. Five per cent of the supernatant from each pelleted sample was analysed simultaneously to estimate the per cent of total phytochrome bound. For quantification of radioactive protein, bands were excised and counted in a scintillation counter.

Received 22 March; accepted 14 June 1999.

- Kendrick, R. E. & Kronenberg, G. H. M. *Photomorphogenesis in Plants* 2nd edn (Kluwer Academic, Dordrecht, Netherlands, 1994).
- Quail, P. H. *et al.* Phytochromes: Photosensory perception and signal transduction. *Science* **268**, 675–680 (1995).
- Quail, P. H. The phytochrome family: Dissection of functional roles and signalling pathways among family members. *Phil. Trans. R. Soc. Lond. B* **353**, 1399–1403 (1998).
- Deng, X.-W. & Quail, P. H. Signaling in light-controlled development. *Semin. Cell Dev. Biol.* (in the press).
- Ni, M., Tepperman, J. M. & Quail, P. H. PIF3, a phytochrome-interacting factor necessary for normal photoinduced signal transduction, is a novel basic helix–loop–helix protein. *Cell* **95**, 657–667 (1998).
- Mathews, S. & Sharrock, R. A. Phytochrome gene diversity. *Plant Cell Environ.* **20**, 666–671 (1997).
- Quail, P. H. An emerging molecular map of the phytochromes. *Plant Cell Environ.* **20**, 657–665 (1997).
- Terzaghi, W. B. & Cashmore, A. R. Light-regulated transcription. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **46**, 445–474 (1995).
- Tobin, E. & Kehoe, D. M. Phytochrome regulated gene expression. *Semin. Cell Biol.* **5**, 335–346 (1994).
- Chory, J. *et al.* From seed germination to flowering, light controls plant development via the pigment phytochrome. *Proc. Natl Acad. Sci. USA* **93**, 12066–12071 (1996).
- Wei, N. & Deng, X.-W. The role of the *COP/DET/FUS* genes in light control of *Arabidopsis* seedling development. *Plant Physiol.* **112**, 871–878 (1996).
- Wagner, D. & Quail, P. H. Mutational analysis of phytochrome B identifies a small COOH-terminal domain region critical for regulatory activity. *Proc. Natl Acad. Sci. USA* **92**, 8596–8600 (1995).
- Xu, Y., Parks, B. M., Short, T. W. & Quail, P. H. Missense mutations define a restricted segment in the COOH-terminal domain of phytochrome A critical to its regulatory activity. *Plant Cell* **7**, 1433–1443 (1995).
- Sakamoto, K. & Nagatani, A. Nuclear localization activity of phytochrome B. *Plant J.* **10**, 859–868 (1996).

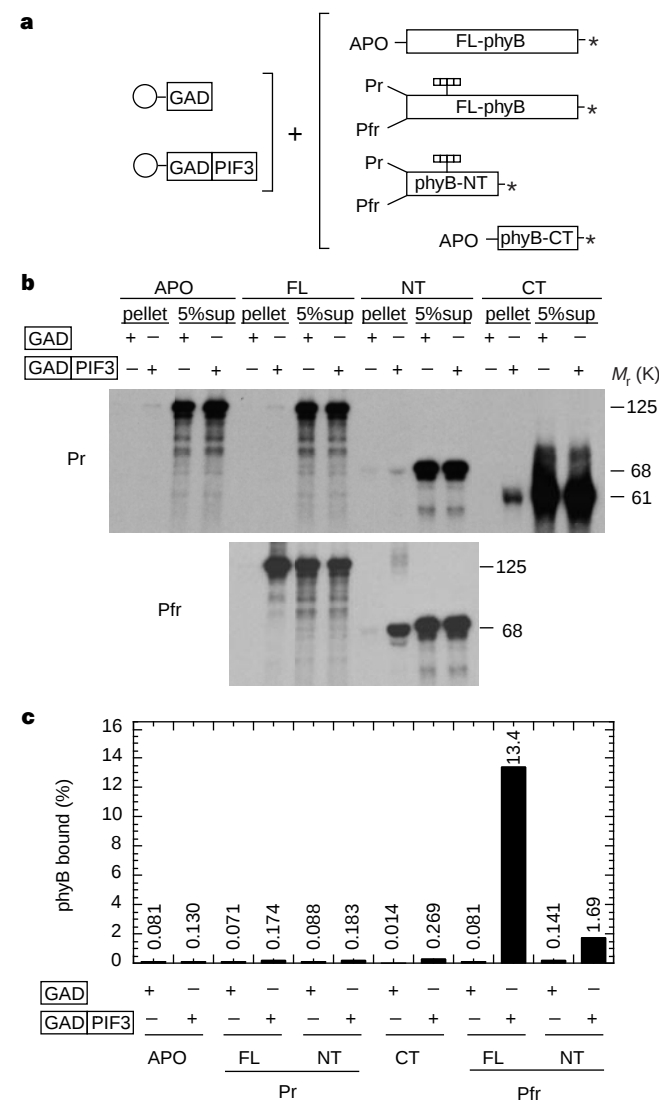


Figure 4 N- and C-terminal domains of phyB are both involved in PIF3 binding. **a**, Experimental design. Recombinant Gal4 activation-domain (GAD) or GAD–PIF3 protein was immobilized on beads, mixed individually with radioactive full-length (FL) or C-terminal-domain (CT) phyB apoproteins (APO), or with chromophore-irradiated, FL or N-terminal-domain (NT) phyB in the Pr or Pfr forms, and incubated for 2 h at 4 °C in darkness before centrifugation. **b**, Autoradiography of FL phyB apoprotein, FL phyB, N-terminal and C-terminal domains of phyB in the entire pellet or 5% of supernatant (5% sup) following incubation as Pr or Pfr in the presence (+) or absence (–) of either GAD or GAD–PIF3. Exposure time for CT sample was 1.5 times longer than for the other samples. **c**, Quantification of PIF3-bound, radioactive phyB shown in **b**, expressed as a per cent of total in supernatant and pellet.

15. Yeh, K.-C. & Lagarias, J. C. Eukaryotic phytochromes: light-regulated serine/threonine protein kinases with histidine kinase ancestry. *Proc. Natl Acad. Sci. USA* **95**, 13976–13981 (1998).
16. Somers, D. E., Sharrock, R. A., Tepperman, J. M. & Quail, P. H. The *hy3* long hypocotyl mutant of *Arabidopsis* is deficient in phytochrome B. *Plant Cell* **3**, 1263–1274 (1991).
17. Scheer, H. in *Techniques in Photomorphogenesis* (eds Smith, H. & Holmes, M. G.) 227–256 (Academic, London, New York, 1984).

Acknowledgements. We thank C. Fairchild for the gift of phycocyanobilin; C. Fairchild, M. Hudson, E. Huq, J. Martinez and E. Monte for critical reading of the manuscript; lab members for discussion and support; and R. Wells for help with the manuscript. This work was supported by grants from the DOG and USDA.

Correspondence and requests for materials should be addressed to P.H.Q. (e-mail: quail@nature.berkeley.edu).

Nucleosome mobilization catalysed by the yeast SWI/SNF complex

Iestyn Whitehouse*, Andrew Flaus*, Bradley R. Cairns†, Malcolm F. White‡, Jerry L. Workman§ & Tom Owen-Hughes*

* Division of Gene Regulation, The Wellcome Trust Building, ‡ Department of Biochemistry, University of Dundee, Dundee DD1 5EH, UK

† Huntsman Cancer Institute, 546 Chipeta Way, Salt Lake City, Utah 84108, USA

§ Howard Hughes Medical Institute, Department of Biochemistry and Molecular Biology, The Pennsylvania State University, University Park, Pennsylvania 16802, USA

The generation of a local chromatin topology conducive to transcription is a key step in gene regulation¹. The yeast SWI/SNF complex is the founding member of a family of ATP-dependent remodelling activities capable of altering chromatin structure both *in vitro* and *in vivo*². Despite its importance, the pathway by which the SWI/SNF complex disrupts chromatin structure is unknown. Here we use a model system to demonstrate that the yeast SWI/SNF complex can reposition nucleosomes in an ATP-dependent reaction that favours attachment of the histone octamer to an acceptor site on the same molecule of DNA (in *cis*). We show that SWI/SNF-mediated displacement of the histone octamer is effectively blocked by a barrier introduced into the DNA, suggesting that this redistribution involves sliding or tracking of nucleosomes along DNA, and that it is achieved by a catalytic mechanism. We conclude that SWI/SNF catalyses the redistribution of nucleosomes along DNA in *cis*, which may represent a general mechanism by which ATP-dependent chromatin remodelling occurs.

To investigate the possibility that the yeast SWI/SNF complex alters the position of nucleosomes on DNA, we generated a model template by ligating nucleosomal cores assembled onto radiolabelled 189-base-pair (bp) DNA fragments to one end of a 1,133-bp piece of naked DNA and separated the ligation products from unligated nucleosomes using streptavidin-coated magnetic beads (Fig. 1a). After washing the beads, barely any radiolabelled DNA was released in the absence of digestion by restriction enzyme (Fig. 1b, lane 1). Digestion of the template with *NheI* released the terminal 180 bp of the construct and electrophoresis on native gels revealed that most of the DNA released had the same mobility as nucleosomal DNA (Fig. 1b, lane 2). If the template was incubated with SWI/SNF complex in the presence of ATP before restriction-enzyme digestion (Fig. 1b, lane 6), most of the radiolabel was released as free DNA, indicating that the SWI/SNF complex promotes the removal of histones from this region of DNA. Consistent with previous studies², this ability of SWI/SNF to displace histones under these conditions was dependent on ATP (Fig. 1b, lane 5).

The SWI/SNF complex could remove histones from DNA by nucleosome disassembly, or by transferring histone octamers from

one piece of DNA to another either by tracking along DNA or by a dissociative pathway³. To distinguish between these possibilities, we digested the template with restriction enzyme before incubating it with SWI/SNF complex. In this case, octamer removal was greatly reduced (Fig. 1b, lane 4). As these reactions in which the DNA was cut before incubation with SWI/SNF were otherwise identical to those performed on the intact template, SWI/SNF-mediated octamer displacement displayed a preference for the attachment of acceptor DNA in *cis*.

To determine the fate of nucleosomes displaced by the SWI/SNF complex, we generated a template containing restriction-enzyme digestion sites at positions that were successively more distant from the original position of the nucleosome (Fig. 2a). Figure 2b (lanes 2, 5, 8 and 11) shows that two distinct species were produced by digestion of this template: the mobility of the faster bands is consistent with release of the appropriate lengths of free DNA from the magnetic beads, the intensity and position of the slower band suggests that this piece of DNA was associated with a single histone octamer, which we confirmed after removing the histone octamer from its DNA by washing the beads with 2 M NaCl before restriction-enzyme digestion. This gave a single band with the same mobility as the faster band (Fig. 2b, lanes 1, 4, 7 and 10). Having established a way to monitor the position of histone octamer on the construct, we investigated the fate of nucleosomes after incubating them with SWI/SNF. We found that 45% of nucleosomes were displaced from the terminal 181 bp of template, but that when the template was digested with the restriction enzymes *BglII*, *NcoI* and *EcoRI*, each of which cuts further along the template, only 34, 18 and 4% of the histone octamers, respectively, were displaced (Fig. 2b, lanes 3, 6, 9 and 12). This ability of SWI/SNF to displace histone octamer from the end of the template but not from longer fragments indicates that it must have been redistributed away from the distal region and along the remainder of the fragment.

It has been inferred from changes in nucleosome positioning that nucleosomes move along DNA by sliding or tracking^{4–9}, but this movement could also result from dissociation of the octamer from

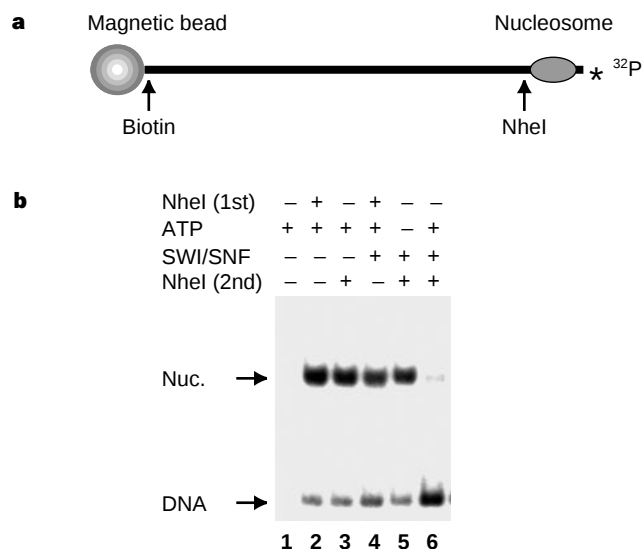


Figure 1 SWI/SNF displaces histones from DNA. **a**, Diagram of the substrate used in the sliding assay. **b**, The template does not enter the matrix of a polyacrylamide gel (lane 1) unless it has been digested with *NheI*, which releases the predominantly nucleosomal 3' end (lane 2). If the template is incubated with SWI/SNF before cutting, most of the ³²P is released as free DNA (lane 6) in an ATP-dependent reaction (lane 5). In otherwise identical reactions, in which the substrate is cut before incubation with SWI/SNF, there is minimal histone displacement (lane 4). Nuc, *NheI*-digested nucleosomal DNA.

its DNA³. The redistribution of nucleosome positioning that we describe, together with the requirement for the attachment of the donor and acceptor DNA in *cis*, strongly support a mechanism for SWI/SNF action that involves nucleosome sliding; however, this is not conclusive as the local concentration of DNA covalently attached to a nucleosome will be high compared with that of the same amount of DNA in free solution and so may bias potential dissociative mechanisms in favour of DNA attached in *cis*. Given that yeast SWI/SNF¹⁰ and RSC¹¹ complexes can both displace histones in *trans*, it is important to distinguish between these pathways.

We therefore introduced a barrier in the DNA to interfere with sliding while maintaining the local concentration of acceptor DNA. As the migration of a Holliday junction in DNA is blocked by a nucleosome¹² and the reverse might also be true, we made constructs containing either a four-way junction (4WJ) or a control DNA sequence which lacked the two arms, between donor and acceptor DNAs (Figs 3a, b). Octamer displacement was largely unaffected by inserting a control linker between the nucleosome and acceptor DNA (compare octamer transfer in Fig. 3c, lanes 1 and 2), but was significantly reduced if inserted with 4WJ DNA (Fig. 3c, lanes 4 and 5). As SWI/SNF can bind 4WJ DNA¹³, we tested the effect of adding 4WJ DNA in *trans* and found that the action of SWI/SNF was not inhibited by 4WJ DNA in *trans* at concentrations 50-fold higher than that needed to prevent sliding when inserted in *cis* (Fig. 3c, lanes 3 and 5). As inserting a barrier between donor and acceptor DNAs disrupts the transfer of octamer to a new site, we conclude that nucleosomes are redistributed by sliding or tracking along DNA.

How can our observations be reconciled with reports that SWI/SNF and related remodelling activities can displace histones in *trans*^{10,11}? We have also detected low levels of octamer displacement

by transfer (Fig. 1b, lane 4) and have found that by increasing the stoichiometry of SWI/SNF to chromatin in the reaction (by decreasing the chromatin concentration) we could promote octamer displacement in *trans* (Fig. 4c, lane 2).

We compared the stoichiometry of SWI/SNF to octamer required for displacement reactions in *cis* and in *trans*. To determine the concentration of SWI/SNF in our preparation, we expressed one subunit of the complex, ARP9 (ref. 14), in *Escherichia coli*, purified it and determined its concentration by spectrophotometric titration against BSA standards (Fig. 4a); we then used known quantities of ARP9 for quantitative western blotting to determine the concentration of our SWI/SNF stock (Fig. 4b). Octamer-displacement reactions were performed with constant SWI/SNF and increasing amounts of unlabelled HeLa oligonucleosomes (Fig. 4c). The results show that yeast SWI/SNF complex can achieve 50% nucleosome displacement in the presence of a 200-fold molar excess of HeLa nucleosomes in 30 min at 30 °C. In contrast, the stoichiometry required for 50% displacement in *trans* was 10-fold higher.

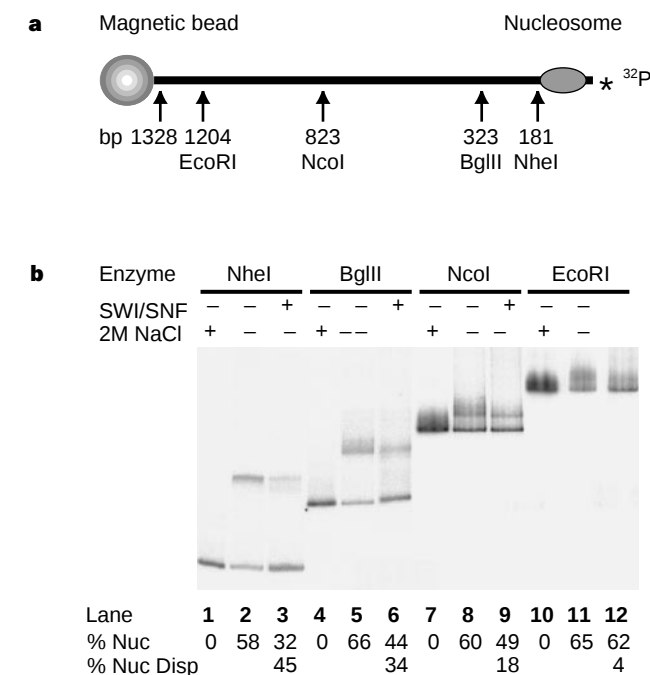


Figure 2 Octamer displaced by SWI/SNF is redistributed along the template. **a**, Restriction-enzyme cutting sites used to study the fate of the histone octamer. **b**, SWI/SNF displaces octamer more efficiently from the distal region of the template than from larger fragments derived from the same template. This is consistent with SWI/SNF causing nucleosomes at the distal end of the construct to become redistributed along its length. The proportion of DNA released as nucleosomal rather than free DNA is indicated below each lane (%Nuc). The proportion of nucleosomes displaced in the presence of SWI/SNF complex is also indicated (% Nuc Disp).

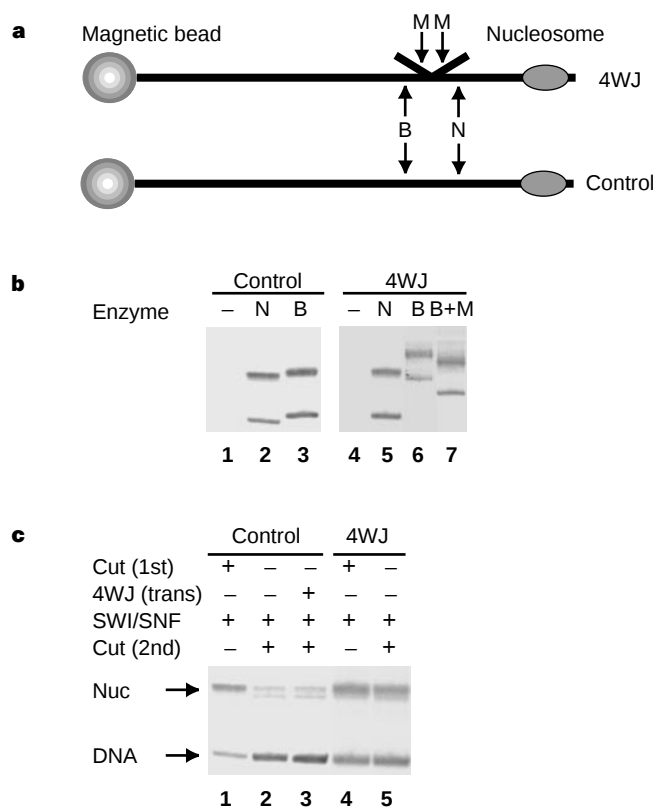


Figure 3 SWI/SNF-mediated transfer of octamer is blocked by a barrier in *cis*. **a**, Constructs containing four-way junction (4WJ) or control DNA inserts. The 25-bp arms of the 4WJ are based on the previously characterized junction 3 (ref. 29) which stack in the orientation indicated in the presence of 3 mM Mg²⁺. Restriction-enzyme cutting sites: B, *BanI*; N, *NheI*; M, *MboI*. **b**, Integrity of these substrates, as confirmed by restriction-enzyme digestion. The fragments released upon digestion with *BanI* ran more slowly than those after *NheI* digestion, and are consistent with the appropriate linker being joined to either nucleosomal or naked DNA (lanes 2, 3 and 5, 6). The presence of the 4WJ arms was confirmed by combined digestion with *BanI* and *MboI*, which removes the distal 18 bp from each arm (lanes 6 and 7). **c**, The constructs containing 4WJ or control DNA linkers were digested with *NheI* before or after remodelling with SWI/SNF. Cutting the control construct after reaction with SWI/SNF and ATP resulted in the proportion of DNA released from the construct increasing from 38% (lane 1) to 63% (lane 2). In contrast, DNA released from the construct containing the 4WJ increased from 47% (lane 4) to 54% (lane 5). Thus, the 4WJ blocked the displacement of octamer in *cis*. Inclusion of 4WJ DNA in *trans* during the reaction on the control template did not inhibit octamer displacement (lane 3). Nuc, digested nucleosomal DNA.

Although stoichiometries strictly should not be compared for chromatin assembled in different reactions, namely by HeLa oligonucleosomes or by core histones^{11,15–17}, the catalytic activity of SWI/SNF in *cis* displacement is significantly greater than observed previously in comparable systems^{10,18,19}. This indicates that the SWI/SNF complex functions more efficiently in octamer displacement when acceptor DNA is attached in *cis*.

The structure of the nucleosome at 2.8 Å resolution indicates that altering the position of a nucleosome involves breaking over 100 interactions between DNA and its histone octamer²⁰. Although these interactions are unlikely all to be disrupted simultaneously, several pathways might involve altering a subset of these interactions at any one time^{4,6,9,21}. A modification of the looping mechanism by which RNA polymerase manages to transcribe through nucleosomes²¹ is attractive in that it might accommodate a displacement of histone octamer both in *cis* and in *trans*. *In vivo*, the balance between displacement in *cis* and in *trans* could be affected by recruitment of high local concentrations of SWI/SNF to regulatory elements²², or by the binding of transcription factors and a restriction of nucleosome mobility¹⁰.

The function of other ATP-dependent chromatin remodelling activities may also involve sliding^{8,23}. Nucleosome sliding, or the consequences of ATP-driven nucleosome sliding on short DNA fragments that restrict nucleosome mobility, may explain observations such as altered nuclease sensitivity of nucleosomal DNA^{15,18,19,24,25}, formation of dinucleosome-like nucleosomes^{16,17}, and changes in nucleosome spacing^{23,26}. Nucleosome sliding may therefore represent a general mechanism by which ATP-dependent chromatin remodelling can activate²⁷ or repress²⁸ transcription. □

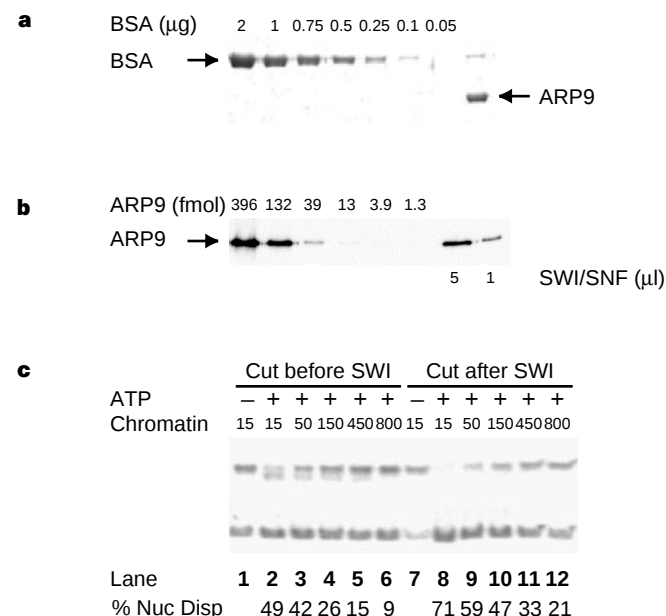


Figure 4 Comparison of the stoichiometry required for SWI/SNF action in *cis* and in *trans*. **a**, Concentration of a preparation of recombinant ARP9, as determined by comparison of the intensity of staining with standard amounts of BSA in a known volume. **b**, Known amounts of ARP9 were used to calculate the amount of SWI/SNF in our preparation by quantitative western blotting. **c**, Sliding was assayed by using constant concentrations of SWI/SNF (0.75 nM) and sliding template, but increasing concentrations (in nM) of chromatin. The template was cut with *NheI* either before (lanes 1–6) or after (lanes 7–12) incubation with the SWI/SNF complex. The concentration of chromatin required for 50% sliding was found to be 15 nM on templates that were cut before SWI/SNF action, and 140 nM on templates that were cut after SWI/SNF action. Thus, SWI/SNF achieved 50% nucleosome displacement at 9.3 (±1.125)-fold lower stoichiometries on templates where acceptor DNA was attached in *cis*.

Methods

Template preparation. For Figs 1 and 4, the large fragment of acceptor DNA was generated by digestion of pCL3 (ref. 15) with *EagI* and filled in with biotin-labelled dCTP (BRL) using Klenow followed by digestion with *AlwNI*. For Figs 2 and 3, acceptor DNA was generated by PCR from a full-length p300 cDNA using the biotin-5'-tagged primer Biotin-5'-cagcgtctgttctacccc and aggggtctt-ggttcggtatgg. The product was cut with the enzyme *BamI* and the 1,137-bp fragment was gel-purified. The shorter fragment was generated by PCR using ³²P-labelled 5'-tcgacaatctttttgtgtc and 5'-ctgaggcaccaggtctgtagctgc as primers and pCC3 as template. The product of this PCR was digested with *AlwNI* or *BamI*, gel-purified and assembled with nucleosomes by transfer from HeLa oligonucleosomes (average length, 2 kb) as described¹⁰. For 3-way ligations, the nucleosomal *AlwNI* fragment was ligated to linker-DNA cassettes generated by annealing the 5'-phosphorylated oligonucleotide 4w11 (gcaccgggagtggtcttagcaagggttggtaccggtatctcaggat), 4w12 (atcctgaagatccgg-taccagctgagcgggttggtatcttcctgctgc), 4w13 (gcacgaagatcaaccacgctcaactcagtcagctagcaggat), 4w14 (ctgtagactgagctgagctctgtaagaccatcccg); 4WJ or 4w11, with a fifth oligonucleotide (ctgaagatccggtaccagcccttctgaagaccatcccg); control, and then purified by Superose-6 gel filtration. Ligations were performed using equimolar amounts (5 pmol) of acceptor DNA, linker where appropriate, and radiolabelled nucleosomes by using 400 U T4 DNA ligase. Unligated nucleosomes were then removed by attachment to 1 mg Dynabeads (Dyna) at room temperature for 1 hour. After attachment, beads were washed 4 times with 100 μl 0.5 M NaCl, 10 mM Tris, pH 7.5, 20 μg ml⁻¹ HeLa oligonucleosomes before storing at 4 °C in 50 μl 0.1 M NaCl, 50 mM HEPES, pH 7.9, 1 mM EDTA, 5 mM DTT, 0.5 mM PMSF, 0.5% N-P40, 10% glycerol, 50 μg ml⁻¹ BSA, 50 μg ml⁻¹ HeLa oligonucleosomes.

Nucleosome sliding assays. Reactions (20 μl) were done in 10 mM HEPES, pH 7.9, 50 mM NaCl, 3 mM MgCl₂, 1 mM Mg-ATP, 5% glycerol, 1 mM PMSF, 0.1 mM DTT, containing 2.5 μg ml⁻¹ total chromatin and 50 fmol template unless stated otherwise. Templates were incubated with or without restriction enzyme (5U) before addition of SWI/SNF¹⁸ for 30 min at 30 °C, then further incubated for 30 min at 37 °C with or without restriction enzyme. All incubations involving immobilized templates were done with mixing or rotation. Nucleosomes were separated from DNA by 4% PAGE, following competition with nucleosomes to remove SWI/SNF as described¹⁹.

SWI/SNF quantification. Quantification was performed using AIDA software and FUJI FLA100 and FLA2000 imaging systems. When determining the concentration of ARP9, SYPRO orange (Molecular Probes) was used according to standard protocols. Detection was with FLA2000 and the concentration of ARP9 determined from a standard curve. The ARP9 in Fig. 4a represents 1 μl stock and was calculated to contain 0.66 μg ARP9. Determination by Bradford assay and Coomassie-blue staining were in agreement, giving values of 0.5 and 0.7 μg, respectively. For quantitative western blotting, ARP 9 antiserum was generated as described¹⁴. A signal generated using horseradish peroxidase-linked secondary antibody and SuperSignal chemiluminescent substrate (Pierce) was detected with a CCD camera and the amount of SWI/SNF determined from a standard curve of band intensity against fmol ARP9.

Received 2 June; accepted 30 June 1999.

- Workman, J. L. & Kingston, R. E. Alteration of nucleosome structure as a mechanism of transcriptional regulation. *Annu. Rev. Biochem.* **67**, 545–579 (1998).
- Peterson, C. L. & Tamkun, J. W. The SWI–SNF complex: a chromatin remodeling machine? *Trends Biochem. Sci.* **20**, 143–146 (1995).
- Steger, D. J. & Workman, J. L. Remodeling chromatin structures for transcription: what happens to the histones? *Bioessays* **18**, 875–884 (1996).
- Flaus, A. & Richmond, T. J. Positioning and stability of nucleosomes on MMTV 3'LTR sequences. *J. Mol. Biol.* **275**, 427–441 (1998).
- Beard, P. Mobility of histones on the chromatin of simian virus 40. *Cell* **15**, 955–967 (1978).
- van Holde, K. E., Yager, T. D. in *Structure and Function of the Genetic Apparatus* (ed. Nicolini, C. T.) 35–53 (Plenum, New York, 1985). **one ed. OK?**
- Pennings, S., Meersseman, G. & Bradbury, E. M. Mobility of positioned nucleosomes on 5S rDNA. *J. Mol. Biol.* **220**, 101–110 (1991).
- Pazin, M. J., Bhargava, P., Geiduschek, E. P. & Kadonaga, J. T. Nucleosome mobility and the maintenance of nucleosome positioning. *Science* **276**, 809–812 (1997).
- Varga-Weisz, P. D. & Becker, P. B. Chromatin-remodeling factors: machines that regulate? *Curr. Opin. Cell Biol.* **10**, 346–353 (1998).
- Owen-Hughes, T., Uley, R. T., Cole, J., Peterson, C. L. & Workman, J. L. Persistent site-specific remodeling of a nucleosome array by transient action of the SWI/SNF complex. *Science* **273**, 513–516 (1996).
- Lorch, Y., Zhang, M. & Kornberg, R. D. Histone octamer transfer by a chromatin-remodelling complex. *Cell* **96**, 389–392 (1999).

12. Grigoriev, M. & Hsieh, P. A histone octamer blocks branch migration of a Holliday junction. *Mol. Cell Biol.* **17**, 7139–7150 (1997).
13. Quinn, J., Fyrborg, A. M., Ganster, R. W., Schmidt, M. C. & Peterson, C. L. DNA-binding properties of the yeast SWI/SNF complex. *Nature* **379**, 844–847 (1996).
14. Cairns, B. R., Erdjument-Bromage, H., Tempst, P., Winston, F. & Kornberg, R. D. Two actin-related proteins are shared functional components of the chromatin-remodeling complexes RSC and SWI/SNF. *Mol. Cell* **2**, 639–651 (1998).
15. Logie, C. & Peterson, C. L. Catalytic activity of the yeast SWI/SNF complex on reconstituted nucleosome arrays. *EMBO J.* **16**, 6772–6782 (1997).
16. Lorch, Y., Cairns, B. R., Zhang, M. & Kornberg, R. D. Activated RSC–nucleosome complex and persistently altered form of the nucleosome. *Cell* **94**, 29–34 (1998).
17. Schnitzler, G., Sif, S. & Kingston, R. E. Human SWI/SNF interconverts a nucleosome between its base state and a stable remodeled state. *Cell* **94**, 17–27 (1998).
18. Côté, J., Quinn, J., Workman, J. L. & Peterson, C. L. Stimulation of GAL4 derivative binding to nucleosomal DNA by the yeast SWI/SNF complex. *Science* **265**, 53–60 (1994).
19. Côté, J., Peterson, C. L. & Workman, J. L. Perturbation of nucleosome core structure by the SWI/SNF complex persists after its detachment, enhancing subsequent transcription factor binding. *Proc. Natl Acad. Sci. USA* **95**, 4947–4952 (1998).
20. Luger, K., Mader, A. W., Richmond, R. K., Sargent, D. F. & Richmond, T. J. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* **389**, 251–260 (1997).
21. Studitsky, V. M., Clark, D. J. & Felsenfeld, G. Overcoming a nucleosomal barrier to transcription. *Cell* **83**, 19–27 (1995).
22. Fryer, C. J. & Archer, T. K. Chromatin remodelling by the glucocorticoid receptor requires the BRG1 complex. *Nature* **393**, 88–91 (1998).
23. Varga-Weisz, P. D., Blank, T. A. & Becker, P. Energy-dependent chromatin accessibility and nucleosome mobility in a cell-free system. *EMBO J.* **14**, 2209–2216 (1995).
24. Varga-Weisz, P. D. et al. Chromatin-remodelling factor CHRAC contains the ATPases ISWI and topoisomerase II. *Nature* **388**, 598–602 (1997); erratum, *ibid.* **389**, 1003 (1997).
25. Tsukiyama, T. & Wu, C. Purification and properties of an ATP dependent nucleosome remodeling factor. *Cell* **83**, 1011–1020 (1995).
26. Ito, T., Bulger, M., Pazin, M. J., Kobayashi, R. & Kadonaga, J. T. ACF, an ISWI-containing and ATP-utilizing chromatin assembly and remodeling factor. *Cell* **90**, 145–155 (1997).
27. Holstege, F. C. et al. Dissecting the regulatory circuitry of a eukaryotic genome. *Cell* **95**, 717–728 (1998).
28. Murphy, D. J., Hardy, S. & Engel, D. A. Human SWI–SNF component BRG1 represses transcription of the *c-fos* gene. *Mol. Cell Biol.* **19**, 2724–2733 (1999).
29. Duckett, D. R. et al. The structure of the Holliday junction, and its resolution. *Cell* **55**, 79–89 (1988).

Acknowledgements. We thank C. Peterson for pCL3; and D. Lilley, J. Allan and members of the Division of Gene Regulation, University of Dundee, for valuable discussion and support. This work was supported primarily by a Wellcome Trust Career Development Award to T.O.-H.; A.E. is an EMBO long-term fellow; I.W. is a BBSRC postgraduate student. A generous donation from the Rappaport Inter-maritime Foundation was used to purchase imaging equipment.

Correspondence and requests for materials should be addressed to T.O.-H. (e-mail: townhughes@bad.dundee.ac.uk).

Four-helical-bundle structure of the cytoplasmic domain of a serine chemotaxis receptor

Kyeong Kyu Kim^{*†‡}, Hisao Yokota[†] & Sung-Hou Kim^{*†§}

^{*} Department of Chemistry, [†] Physical Biosciences Division of Lawrence Berkeley National Laboratory, [§] 220 Melvin Calvin Laboratory, University of California, Berkeley, California 94720-5230, USA

The bacterial chemotaxis receptors are transmembrane receptors with a simple signalling pathway which has elements relevant to the general understanding of signal recognition and transduction across membranes, how signals are relayed between molecules in a pathway, and how adaptation to a persistent signal is achieved¹. In contrast to many mammalian receptors which signal by oligomerizing upon ligand binding², the chemotaxis receptors are dimeric even in the absence of their ligands, and their signalling does not depend on a monomer–dimer equilibrium³. Bacterial chemotaxis receptors are composed of a ligand-binding domain, a transmembrane domain consisting of two helices TM1 and TM2, and a cytoplasmic domain. All known bacterial chemotaxis receptors have a highly conserved cytoplasmic domain, which unites signals from different ligand domains into a single signalling pathway to flagella motors. Here we report the crystal structure of the cytoplasmic domain of a serine chemotaxis receptor of *Escherichia coli*, which reveals a 200 Å-long coiled-coil of two

antiparallel helices connected by a ‘U-turn’. Two of these domains form a long, supercoiled, four-helical bundle in the cytoplasmic portion of the receptor.

The three-domain structure of bacterial chemotaxis receptors is shown schematically in Fig. 1. The cytoplasmic domain is itself divided into linker, methylation, and signalling regions⁴, and is able to send signals in the same way as the intact receptor under different methylation states^{5–7}. We have cloned and expressed the cytoplasmic domain of the serine receptor of *Escherichia coli* corresponding to the highest methylation state by mutating all four methylation sites to glutamine residues.

The crystal structure of this domain, cTsrQ, reveals that it is dimeric with a partial two-fold non-crystallographic symmetry axis along most of the length of the dimer (Fig. 2). Each monomer is a ~70-turn α -helix folded back on itself to form two long antiparallel coiled α -helices, the N helix and the C helix, connected by a short ‘U-turn’. Thus, a dimer of cTsrQ is a long four-helix supercoiled helical bundle in which four α -helices pack against each other with primarily, but not exclusively, hydrophobic residues inside the bundle. Each of the four helices in the bundle is a continuous coiled-coil with a pitch of about 200 Å, but the curvature of supercoiling of the N and C helices is not identical. This arrangement is similar to earlier models^{4,5}, but the non-helical coils and one β -strand predicted in one of these models⁴ have not been found in the crystal structure. In our model, the ordered amino terminus in the crystal structure starts from residue 294 in one monomer (monomer A) and from residue 300 in the other (monomer B). The ordered N helix (residues 294–389 in monomer A and 300–389 in B) is about 40 residues shorter than the C helix (residues 392–520 in both monomers). The carboxy-terminal ‘overhanging’ helix from one monomer makes an antiparallel coiled-coil with that of another monomer related by a crystallographic two-fold symmetry axis (not the non-crystallographic two-fold axis between two monomers in the dimeric structure) which is approximately perpendicular to the length of the molecule.

Most of the length of each N helix is in contact with two C helices, one from the same monomer and the other from the other

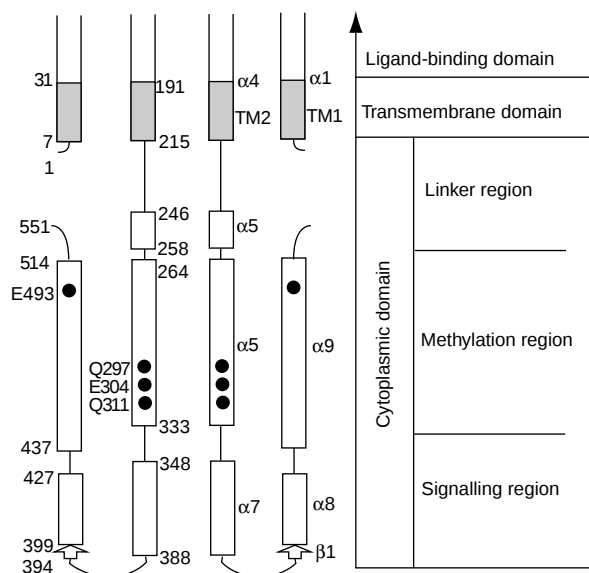


Figure 1 The transmembrane serine chemotaxis receptor of *E. coli*. Assignment of various functional domains along the protein sequence, the secondary-structure assignment, and the predicted topological arrangement⁴ of known functional regions of the cytoplasmic domain are indicated. Predicted helical regions are shown as long rectangles, non-helical coils as thick lines, and a β -strand as an arrow. The first and last residues of each predicted secondary structure are labelled. Predicted transmembrane helices are shaded; methylation sites are shown as filled circles.

[†] Present address: Plant Molecular Biology and Biotechnology Research Center, Department of Molecular Biology, Gyeongsang National University, Chinju 660-701, Korea.

monomer in the same dimer. Similarly, each C helix is in contact with two N helices. As predicted earlier^{4,5}, most of the third and seventh residues starting from residue 302 to the end of the N helix and those starting from the residue 397 to the end of the C helix are buried and can be assigned as 'd' and 'a' positions in the heptad hydrophobic repeat, respectively. However, there is another heptad repeat not noticed previously: starting from residue 305 of the N helix and from residue 396 of the C helix, most of the third and seventh residues are buried. Such 'double heptads' should be a common feature of all multiple-helix bundles, because in such bundles each helix is in contact with two other helices. The relative registration between the N and C helices in the crystal structure is different from that predicted in an earlier model⁴ by about two turns of helix. Most of the residues at the interhelical contacts are hydrophobic, but there are several hydrophilic residues in these contacts which may have caused erroneous secondary-structure or coiled-coil predictions in the earlier model. In our crystal structure, these hydrophilic residues form hydrogen bonds with hydrophilic residues from other helices at the contact surfaces that are sometimes mediated by water molecules. At the present resolution, the helical structure of the backbones is clearly recognizable (Fig. 3), but the side chains of the residues with high temperature factors (residues 286–349 and 431–474) are not well resolved, or are unrecognizable in the electron density map (Fig. 3a). These regions correspond to the amino terminus and to the middle of the two helical coiled-coils of each monomer, and we interpret them to be more flexible and mobile than the rest of the structure (Fig. 2, left). Consequently, the helices in these regions are not as conformationally regular as canonical α -helices: the helical geometry deviates more around residues 294, 295, 330, 391, 450 and 517–520 in monomer A, and around residues 335, 345, 390–392, 439–443 and 519–520 in monomer B.

Of the residues at four methylation sites in monomer B of the cTsrQ dimer, the first (Q297) is disordered and not visible in the electron density map. For the remaining two methylation residues in the N helix, the peptide backbones are ordered and clearly recognizable in the map, but their side-chain atoms have high temperature factors and are not recognizable. The same is true for the other monomer in the dimer, except that the backbone of Q297 in this monomer is also visible in the map. The methylation site in the C helix (Q493), however, is well ordered in both monomers and has a low temperature factor. All methylation-site residues are exposed on the surface of the dimeric structure (Fig. 2). Although the side chains of the three methylation site residues of cTsrQ (Q297, Q304 and Q311) are disordered because of high temperature factors, their predicted (on the basis of their β -carbon positions) side-chain locations in the N helix of one monomer are at hydrogen-bonding distances from the residues of E479, Q472 and E465 of the C helix of the other monomer, respectively.

The protein sequence of the signalling region is highly conserved among all bacterial chemotaxis receptors⁴. This region consists of two antiparallel coiled-coil helices connected by a short 'U turn'. Compared with the methylation region, the signalling region has lower temperature factors and is more well ordered. An interesting feature of this region is that three cTsrQ monomers, one each from three dimers related by a three-fold crystallographic symmetry, form a cluster. Upon formation of trimers, 970 Å² of the surface area is buried, which corresponds to about 4.1% of the total surface area. The clustering occurs as a result of extensive interaction among the residues near the U turn that are completely conserved in all chemotaxis receptors (Fig. 4a, b): Phe 373 of one monomer interacts with Ile 377 of a neighbouring monomer related by three-fold symmetry; Leu 380 and Val 384 of one monomer make hydrophobic interactions with Leu 378 and Val 398, respectively, of the symmetry-equivalent monomer; Asn 376 of one monomer and Arg 409 of the symmetry-related monomer form a hydrogen bond; three symmetrically equivalent Asn 381 residues form a

hydrogen-bonded network; and Glu 385 in one monomer and Arg 388 of the symmetry-related monomer form a salt bridge and a hydrogen bond through water. As these extensive interactions involve residues that are strictly conserved in all chemotaxis receptors, the trimerization of the cTsr dimers may be an intrinsic property of the intact Tsr chemotaxis receptors. Such clustering may be related to the large patches of chemotaxis receptor clusters observed *in vivo*⁸, and to the clustering of signalling domains *in vitro*⁹.

We have constructed a model of the intact *E. coli* serine chemoreceptor (Fig. 5) by combining the crystal structure of cTsrQ and a model of the ligand-binding domain of Tsr based on the crystal structure of the ligand-binding domain of Tar of *E. coli*¹⁰. The receptor model is 380 Å long and consists of an 80 Å-long ligand-binding domain, a 40 Å-long transmembrane domain, and a 260 Å-long cytoplasmic domain. The dimer of the intact receptor makes a long four-helix bundle from the top of the ligand-binding domain to the bottom of the signalling region, apart from a short two-helix coiled-coil (residues 222–241) in the linker region and an additional four, shorter helices in the ligand-binding domain. In the regions of the model where no crystallographic information was used, we found several favourable interactions at the core of the four-helical bundle, in support of our model for these regions.

There are two conflicting hypotheses on the molecular mechanism by which the signal is transmitted through the membrane: in one, the signal is transmitted through a monomer unit^{11–14}, and in the other it is through a dimer unit^{3,5,15,16}. In addition to the two static models described above, the difference in the dynamic properties of the receptors may also account for the signalling transfer mechanism^{16,17}. The crystal structure of cTsrQ and the model of the intact Tsr receptor reveal that most of the length of the receptor comprises a four-helical bundle with a predominantly

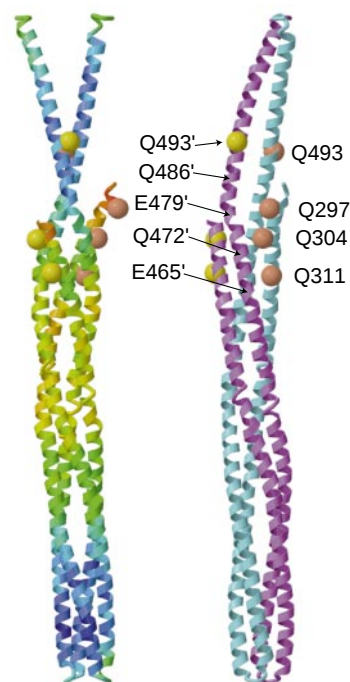


Figure 2 Two views of the cTsrQ dimer structure related by a 90-degree rotation around the non-crystallographic two-fold axis along the length of the molecules. The methylation sites are shown as yellow balls in one monomer and orange balls in the other. Right: one monomer of cTsrQ is shown in purple and the other is light blue; left: residues with high temperature factors are shown in red and those with low temperature factors in blue. The average *B* factor in the well-defined region (residues 361–420) is 36.3 Å² and in the other part is 79.7 Å². The figures were generated by the program MOLSCRIPT²⁰.

Table 1 Structure determination

Data collection					
Crystal	Wavelength (Å)	Resolution (Å)	Completeness (%)	<i>I</i> / σ	<i>R</i> _{sym} (%)*
cTsrQ	1.08	30.0–2.6	95.6	4.1	12.5
cTsrQ(Se-Met)	0.9789 (edge)	50.0–3.5	98.1	20.5	8.9
(selenium)	0.9786 (peak)	50.0–3.5	98.1	20.4	9.0
	0.9488 (remote)	50.0–3.5	98.0	21.9	7.7
Phasing					
	Phasing power†	<i>R</i> _{cullis} ‡	<i>R</i> _{cullis} (ano)		
Edge			0.60		
Peak	0.60	0.70	0.60		
Remote	0.70	0.66	0.70		
Figure of merit: 0.457 (30.0–3.5 Å)					
Refinement					
Resolution	30.0–2.6 Å				
<i>R</i> _{crystal} (<i>R</i> _{free}) in %	24.3 (27.6)				
Completeness (%)	51.6 (75.0)¶				
Unique reflection (2σ cut-off)	21,860				
No. of protein atoms	3,243				
No. of water atoms	162				
Mean <i>B</i> factors (Å²)	64.9				
R.m.s. deviation of bond (Å)	0.011				
R.m.s. deviation of angle (deg)	1.51				

* $R_{\text{sym}} = \sum |I_{\text{obs}} - I_{\text{avg}}| / \sum I_{\text{avg}}$.
† Phasing power is the mean F_n divided by the r.m.s. lack of error.
‡ R_{cullis} is the r.m.s. lack of error divided by the isomorphous difference.
§ R_{free} was calculated with 5% of reflections.
‖ The value in parentheses represents the completeness against the total number of reflections in the resolution range of 30.0 to 2.6 Å along the *c*-axis, and 3.5 Å along the *a* and *b* axes. The completeness of between 30.0 to 3.5 Å for the three axes is 85.0%.

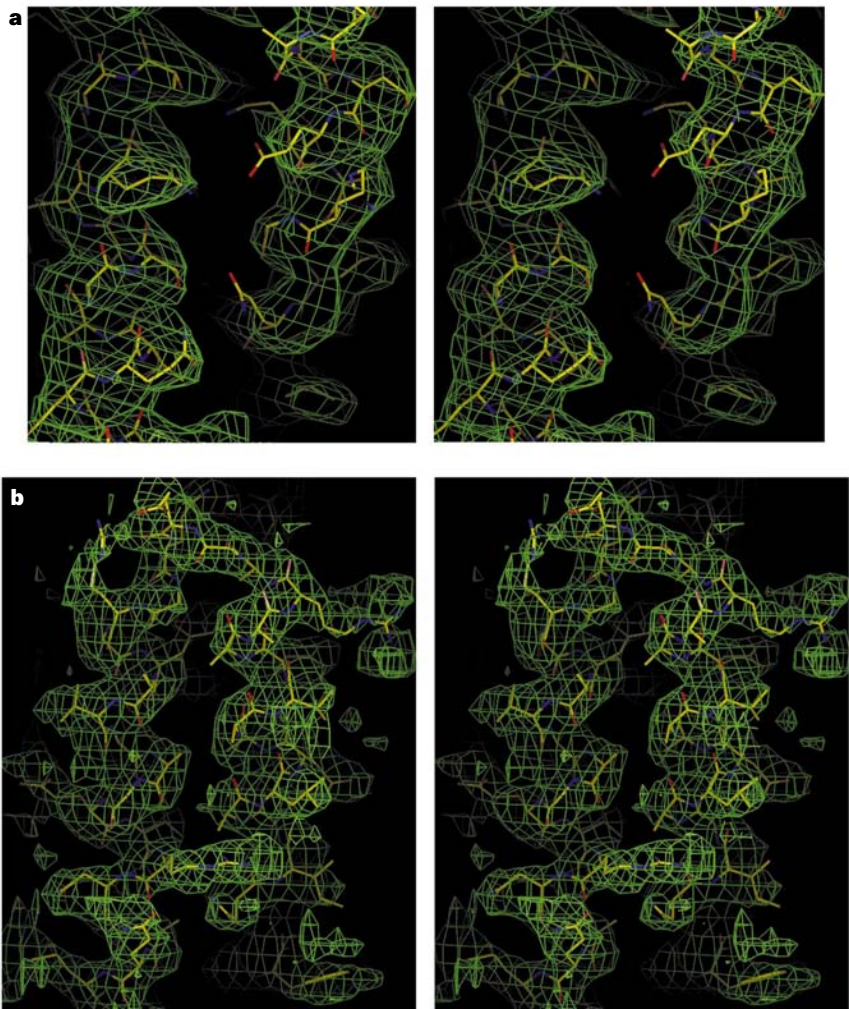


Figure 3 Typical regions of the solvent-flattened experimental electron density map (contoured at a 1.2 σ level) superposed with the final structural model of cTsrQ. **a**, A region with high temperature factors, where some side-chain densities are not observed, although the backbone densities are clear. **b**, The region around the U-turn.

hydrophobic core: this feature suggests that the signalling properties of hybrid dimeric receptors with only one cytoplasmic domain, used previously to support a monomeric mechanism of signalling^{11–14}, may need to be reinterpreted. Such a monomeric cytoplasmic domain would expose many hydrophobic residues normally protected inside the four-helical bundle, so signalling from individual monomers is less likely.

All four glutamine residues at the methylation sites of one cTsrQ monomer are at the positions that could form hydrogen bonds with residues of the other monomer in a dimeric structure of cTsrQ, which corresponds to the cytoplasmic domain of the receptor with the highest kinase activity^{5,18}. This may mean that a receptor dimer having a high kinase activity may be conformationally 'frozen' compared with a dimer with low kinase activity, either because more hydrogen bonds are formed, as in the Q mutant, or because

favourable hydrophobic interactions involving methyl groups are possible, as in a fully methylated wild-type receptor. Thus, the structure indicates that changes in superhelicity and/or dynamic flexibility¹⁶ of the four-helix bundle may provide a mechanism for signalling.

Mutations in either the cytoplasmic domain of intact receptors or in the isolated domains can generate biased signalling phenotypes *in vivo*, and 'locked' or altered kinase activity *in vitro*^{6,7,9,19–22}. In the crystal structure, most of the mutation sites are buried in the four-helical bundle of the cytoplasmic domain, so they may alter the conformation and/or a dynamic property of the four-helical bundle. A few mutation sites associated with altered kinase activity or cellular behaviour are found on the surface of the four-helical bundle of the dimer and may lie on or near the recognition sites for interacting proteins such as CheW and CheA. □

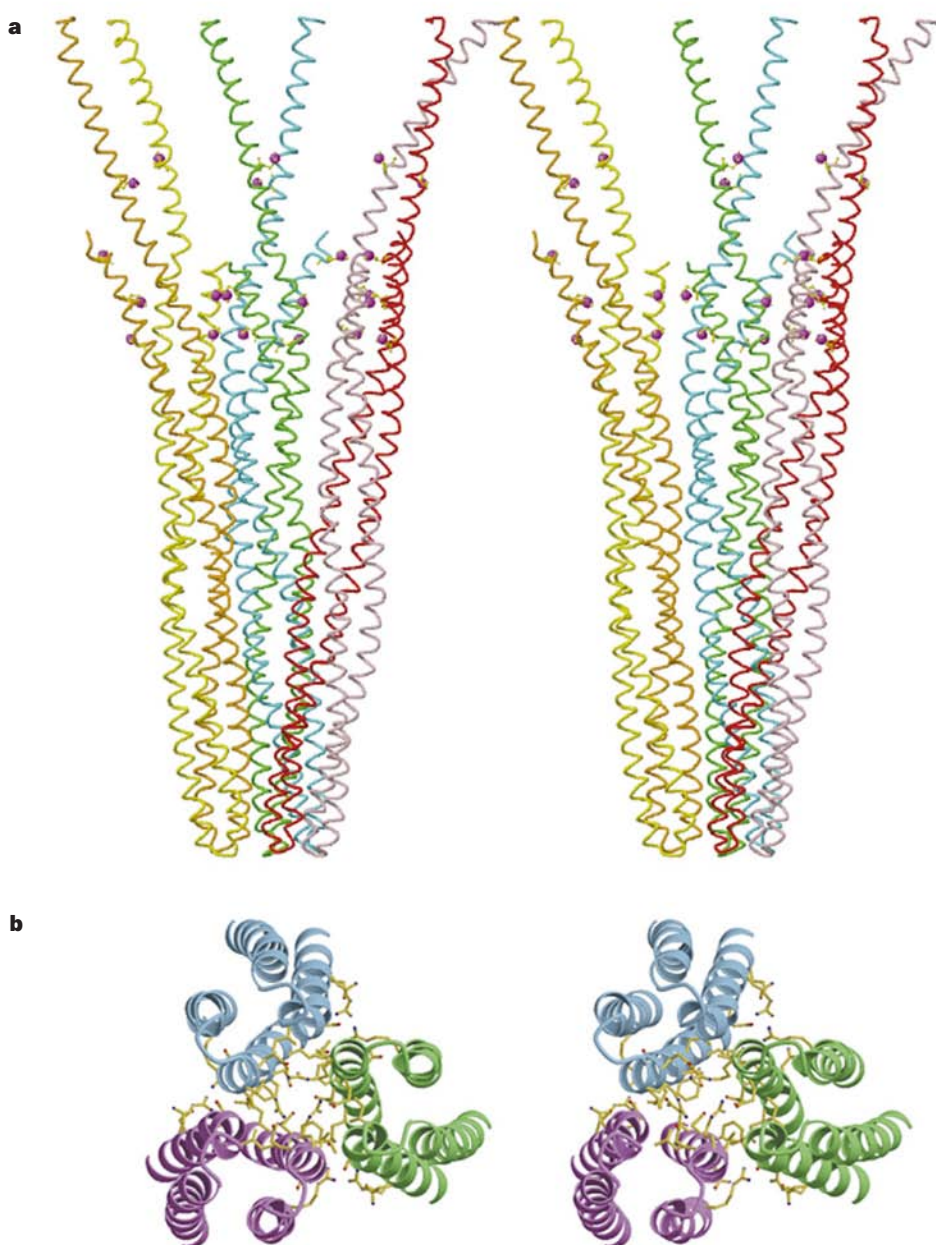


Figure 4 Trimer of cTsr dimers. **a**, A stereo diagram of a trimer of cTsrQ dimers in the crystal. Each monomer is coloured differently. The methylation sites are shown as small balls. **b**, A stereo view of the details of interaction at the contact

site of three cTsrQ monomers, one each from three dimeric cTsrQ molecules. Each dimer is in one colour.

Methods

Cloning, mutagenesis and purification. A DNA insert coding for the *E. coli* Tsr cytoplasmic domain spanning amino acids 286 to 526 was amplified by PCR from plasmid HB915 containing the *E. coli* Tsr gene. The plasmid was mutated to change the wild-type QEQE (amino acids at the four methylation sites) genotype to QQQQ (Q-mutant) by using the Kunkel²³ site-specific mutagenesis protocol; the mutation was verified by DNA sequencing. The cytoplasmic domain of the serine receptor was purified as described⁷, with minor modifications: 5-ml Pharmacia HiTrap Q was used as an anion-exchange column in a buffer of 25 mM bis-Tris, 1 mM EDTA (pH 6.5), and a phenyl-Sepharose hydrophobic-interaction column (BioRad) was used in 25 mM potassium phosphate buffer (pH 7.0), 1 mM EDTA, with a linear gradient of 750 mM to 0.0 mM NaCl.

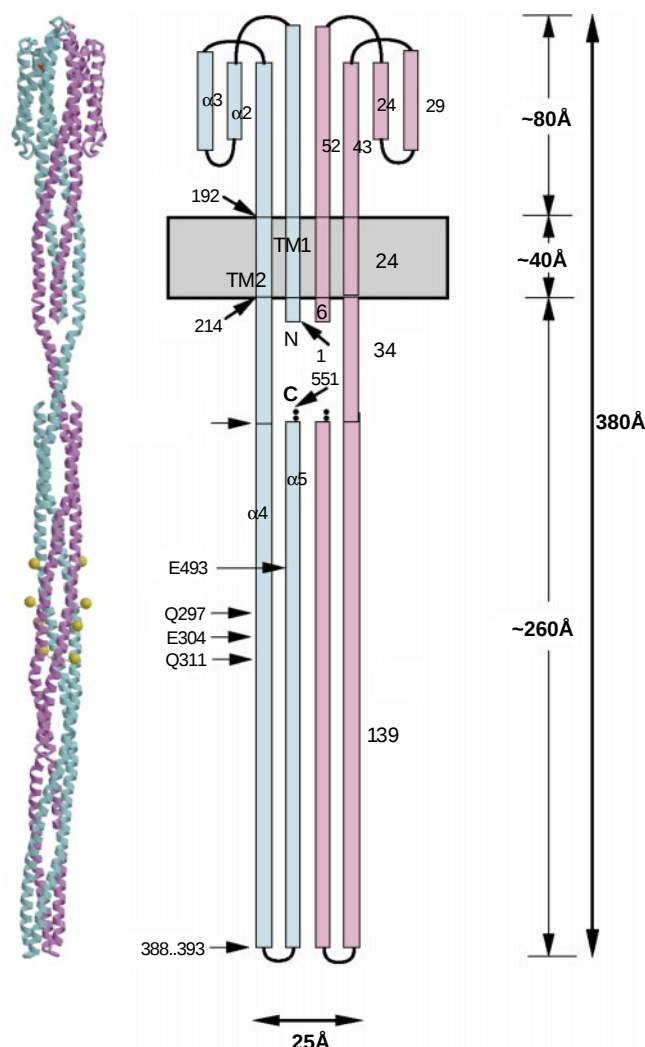


Figure 5 Model of an intact *E. coli* Tsr receptor dimer. Right: diagram of an intact Tsr receptor dimer model, with one monomer in blue and the other in pink. The presumed membrane bilayer is represented by a grey band. Landmark residues are labelled in the smaller font, and the numbers of residues in helical sections are shown in a larger font. The length of each domain is indicated. Left: ribbon diagram of the intact Tsr dimer model viewed perpendicularly to the non-crystallographic two-fold symmetry axis. The dimensions are scaled to match those of the figure at right. One monomer is in purple, and the other in cyan. Methylation sites are marked by yellow balls in one monomer and orange balls in the other, and the ligand serine is drawn as a red ball partially hidden at upper left corner. Computer-modelled portions without crystallographic models are less reliable. They include residue 1 to the end of TM1, TM2 and up to residue 293, and residues 521 to 551, especially the CheR-binding region at the C-terminal end.

Crystallization and data collection. The initial crystallization conditions were found by the sparse matrix screening method²⁴ using a Hampton screen (Hampton Research). The protein solution was prepared in a buffer containing 1 mM EDTA and 25 mM HEPES (pH 7.4) at 5–10 mg ml⁻¹ protein. Crystals of cTsrQ were obtained at 4 °C from a solution containing 0.1 M N-[2-acetamide]-2-iminodiacetic acid (pH 6.5), 15–20% MPD (2-methyl-2,4-pentanediol), 0.2 M sodium citrate. Selenomethionine-substituted protein was crystallized from a solution containing 15–20% MPD, 0.1 M HEPES (pH 7.5), 0.2 M magnesium sulphate. Crystals were transferred to mother liquor containing a higher concentration of MPD (increased by up to 25%) for 24 h before flash-freezing and collecting data. X-ray data from native cTsrQ crystals were collected to a resolution of 2.6 Å and those from the selenomethionine-substituted crystal to 3.5 Å resolution from three wavelengths (Table 1). The distribution of the intensities of diffraction pattern was very anisotropic: diffraction intensities were observed along the *c*-axis up to 2.6 Å, but along *a* and *b* axes only up to 3.5 Å. We interpreted this anisotropy to indicate the molecular transform of a very long structure having its long axis approximately parallel to the crystallographic *c*-axis. The data were processed and integrated by DENZO and scaled by using SCALEPACK²⁵. The space group of the crystals is *P*321, with unit-cell dimensions *a* = *b* = 132.44 Å, *c* = 134.59 Å, with less than 1% unit-cell variation. Assuming a relative molecular mass of 25,314 and one homodimer per asymmetric unit, the volume-to-mass ratio is 7.4 Å³ per dalton, and the solvent content is estimated to be 83%, which is unusually high.

Structure determination and refinement. X-ray diffraction data collected at three different wavelengths were used for conventional multiple isomorphous replacement. The first selenium site was found in the difference Patterson map between the data collected at the Se-absorption edge and a remote wavelength, and its position was refined with the program MLPHARE²⁶. The other seven selenium sites were found in the difference Fourier map by using the phase calculated from the first selenium site. Eight out of 10 selenium atoms in an asymmetric unit were found and their positions were refined. The initial phases were improved by solvent flattening with DM²⁶ using a solvent content of 80% and were extended to 2.6 Å resolution using the native data. After density modification, electron density revealed a four-helical bundle without any breaks in the backbone chain. Amino acids were assigned from residues using program O (ref. 27). All eight selenium sites corresponded to the expected methionine residues in the fitted sequences. The electron densities for both the main chains and the side chains of residues 350–430 and 475–520 in each monomer were readily interpretable (Fig. 3b), but those for the side chains of residues in the other region (residues 300–349 and 431–474) were weak or non-existent, even though the helical backbones were clear (Fig. 3a). The amino termini (residues 286–293 of one monomer and 286–299 of the other) were disordered, and the carboxy termini of both monomers (residues 521–526) were also disordered. Statistics of the data, phasing and structure refinement are shown in Table 1. As the diffraction intensities were anisotropic, reflections in the resolution range of 30.0 to 3.5 Å along the *a*- and *b*-axes, and 2.6 Å along the *c*-axis, were used for refinement. A total of 21,860 reflections were used to refine the positions and *B* factors of 3,405 atoms in the current model. Several cycles of rigid-body refinement, positional refinement, simulated annealing, and *B*-factor refinement using X-PLOR²⁸ and CNS (A. T. Brünger, personal communication) with the bulk solvent correction, followed by an anisotropic *B*-factor refinement gave an *R* factor of 24.3% and an *R*_{free} of 27.6% at 2.6 Å resolution (anisotropic *B* factor: *B*₁₁ = -16.3, *B*₁₂ = -28.7, *B*₁₃ = 0.0, *B*₂₂ = -16.3, *B*₂₃ = 0.0, and *B*₃₃ = 32.7). Most (85.2%) of the non-glycine amino acids assumed the most favourable conformation and the remaining 14.8% were in the allowed conformational regions in the Ramachandran plot drawn by using PROCHECK²⁹.

Model building of intact receptor. The intact Tsr model was built by combining the crystal structure of cTsrQ and a model of the ligand domain of Tsr based on the crystal structure of the ligand domain of an *E. coli* aspartate chemotaxis receptor, Tar¹⁰. The carboxy-terminal residues 476–520 and amino-terminal residues from two monomers were modelled to form a four-helix bundle. The four-helix-bundle architecture of cTsrQ was used as a template in making the four-helix-bundle portion of the linker region. The transmembrane-spanning and linker regions were assumed to form contiguous coiled-coils because the linker regions have a heptad repeat of hydrophobic residues similar to the coiled-coil regions of the structure. The model was

energetically optimized by using several cycles of molecular dynamics calculations with X-PLOR²⁸ and two-fold symmetry restraints between the two subunits in the dimer.

Received 5 March; accepted 21 June 1999.

1. Falke, J. J., Bass, R. B., Butler, S. L., Chervitz, S. A. & Danielson, M. A. The two-component signaling pathway of bacterial chemotaxis: a molecular view of signal transduction by receptors, kinases, and adaptation enzymes. *Annu. Rev. Cell Biol. Dev. Biol.* **13**, 457–512 (1997).
2. Ullrich, A. & Schlessinger, J. Signal transduction by receptors with tyrosine kinase activity. *Cell* **61**, 203–212 (1990).
3. Chervitz, S. A. & Falke, J. J. Molecular mechanism of transmembrane signaling by the aspartate receptor: a model. *Proc. Natl Acad. Sci. USA* **93**, 2545–2550 (1996).
4. Le Moual, H. & Koshland, D. E. Jr Molecular evolution of the C-terminal cytoplasmic domain of a superfamily of bacterial receptors involved in taxis. *J. Mol. Biol.* **261**, 568–585 (1996).
5. Cochran, A. G. & Kim, P. S. Imitation of *Escherichia coli* aspartate receptor signaling in engineered dimers of the cytoplasmic domain. *Science* **271**, 1113–1116 (1996).
6. Ames, P., Yu, Y. A. & Parkinson, J. S. Methylation segments are not required for chemotactic signalling by cytoplasmic fragments of Tsr, the methyl-accepting serine chemoreceptor of *Escherichia coli*. *Mol. Microbiol.* **19**, 737–746 (1996).
7. Surette, M. G. & Stock, J. B. Role of α -helical coiled-coil interactions in receptor dimerization, signaling, and adaptation during bacterial chemotaxis. *J. Biol. Chem.* **271**, 17966–17973 (1996).
8. Maddock, J. R. & Shapiro, L. Polar location of the chemoreceptor complex in the *Escherichia coli* cell. *Science* **259**, 1717–1723 (1993).
9. Liu, Y., Levit, M., Lurz, R., Surette, M. G. & Stock, J. B. Receptor-mediated protein kinase activation and the mechanism of transmembrane signaling in bacterial chemotaxis. *EMBO J.* **16**, 7231–7240 (1997).
10. Chi, Y. I., Yokota, H. & Kim, S.-H. Apo structure of the ligand-binding domain of aspartate receptor from *Escherichia coli* and its comparison with ligand-bound or pseudoligand-bound structures. *FEBS Lett.* **414**, 327–332 (1997).
11. Lynch, B. A. & Koshland, D. E. Jr The fifth Datta Lecture. Structural similarities between the aspartate receptor of bacterial chemotaxis and the trp repressor of *E. coli*. Implications for transmembrane signaling. *FEBS Lett.* **307**, 3–9 (1992).
12. Gardina, P. J. & Manson, M. D. Attractant signaling by an aspartate chemoreceptor dimer with a single cytoplasmic domain. *Science* **274**, 425–426 (1996).
13. Tatsuno, I., Homma, M., Oosawa, K. & Kawagishi, I. Signaling by the *Escherichia coli* aspartate chemoreceptor Tar with a single cytoplasmic domain per dimer. *Science* **274**, 423–425 (1996).
14. Hughson, A. G. & Hazelbauer, G. L. Detecting the conformational change of transmembrane signaling in a bacterial chemoreceptor by measuring effects on disulfide cross-linking *in vivo*. *Proc. Natl Acad. Sci. USA* **93**, 11546–11551 (1996).
15. Milburn, M. V. *et al.* Three-dimensional structures of the ligand-binding domain of the bacterial aspartate receptor with and without a ligand. *Science* **254**, 1342–1347 (1991).

16. Kim, S.-H. “Frozen” dynamic dimer model for transmembrane signaling in bacterial chemotaxis receptors. *Protein Sci.* **3**, 159–165 (1994).
17. Maruyama, I. N., Mikawa, Y. G. & Maruyama, H. I. A model for transmembrane signaling by the aspartate receptor based on random-cassette mutagenesis and site-directed disulfide cross-linking. *J. Mol. Biol.* **253**, 530–546 (1995).
18. Borkovich, K. A., Alex, L. A. & Simon, M. I. Attenuation of sensory receptor signaling by covalent modification. *Proc. Natl Acad. Sci. USA* **89**, 6756–6760 (1992).
19. Ames, P. & Parkinson, J. S. Transmembrane signaling by bacterial chemoreceptors: *E. coli* transducers with locked signal output. *Cell* **55**, 817–826 (1988).
20. Ames, P., Chen, J., Wolff, C. & Parkinson, J. S. Structure–function studies of bacterial chemosensors. *Cold Spring Harb. Symp. Quant. Biol.* **53**, 59–65 (1988).
21. Danielson, M. A., Bass, R. B. & Falke, J. J. Cysteine and disulfide scanning reveals a regulatory α -helix in the cytoplasmic domain of the aspartate receptor. *J. Biol. Chem.* **272**, 32878–32888 (1997).
22. Bass, R. B. & Falke, J. J. Detection of a conserved α -helix in the kinase docking region of the aspartate receptor by cysteine and disulfide scanning. *J. Biol. Chem.* **273**, 25006–25014 (1998).
23. Kunkel, T. A. Rapid and efficient site-specific mutagenesis without phenotypic selection. *Proc. Natl Acad. Sci. USA* **82**, 488 (1985).
24. Jancarik, J. & Kim, S.-H. Sparse matrix sampling: a screening method for crystallization of proteins. *J. Appl. Crystallogr.* **24**, 409–411 (1991).
25. Otwinowski, Z. in *Data Collection and Processing* (eds Sawyer, L., Isaacs, N. & Bailey, S.) 56–62 (SERC Daresbury Laboratory, Warrington, UK, 1993).
26. Collaborative Computing Project No. 4. The CCP4 suite: Programs for protein crystallography. *Acta Crystallogr. D* **50**, 760–763 (1994).
27. Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. Improved methods for binding protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A* **47**, 110–119 (1991).
28. Brünger, A. T. *X-PLOR Version 3.1* (Yale Univ. Press, New Haven, CT, 1993).
29. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. Procheck: A program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26**, 283–291 (1993).
30. Kraulis, P. I. MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* **24**, 946–950 (1991).

Acknowledgements. We thank R. Sweet, P. Kuhn and H. Bellamy for data collection; D. King for performing the electrospray mass spectrometry; C. Park for plasmid HB915; K. Kamata for help with sample preparation; Z. Zhang and E. Berry for preparing some of the figures; and J. Falke and S. Parkinson for discussion. This work was supported by grants from the Office of Biological and Environmental Research, Office of Science, DOE and the NIH.

Correspondence and requests for materials should be addressed to S.H.K. at the Melvin Calvin Laboratory (e-mail: shkim@LBL.gov). The accession code for the atomic coordinates of the structure at the Protein Data Bank is 1QU7.

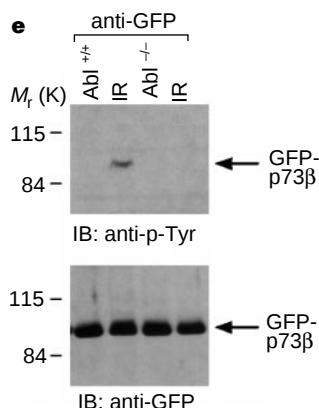
errata

p73 is regulated by tyrosine kinase c-Abl in the apoptotic response to DNA damage

Zhi-Min Yuan, Hisashi Shioya, Takatoshi Ishiko, Xiangao Sun, Jijie Gu, YinYin Huang, Hua Lu, Surender Kharbanda, Ralph Weichselbaum & Donald Kufe

Nature **399**, 814–817 (1999)

Figure 2e as published was a duplicate of Fig. 2d. The correct figure is shown below.

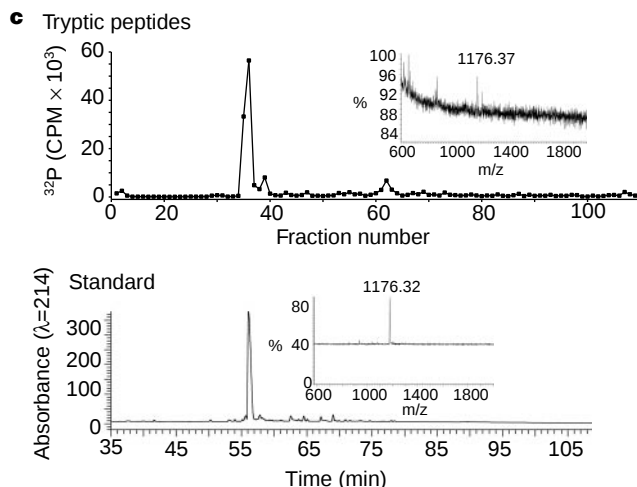


Regulation of endothelium-derived nitric oxide production by the protein kinase Akt

David Fulton, Jean-Philippe Gratton, Timothy J. McCabe, Jason Fontana, Yasushi Fujio, Kenneth Walsh, Thomas F. Franke, Andreas Papapetropoulos & William C. Sessa

Nature **399**, 597–601 (1999)

In Fig. 2c, the lower panel was incorrectly shifted to the right: the figure should have appeared as below. In Fig. 4b, the concentration of free calcium was in micromolar, and not millimolar as published.



Chromatography aids

Detecting, warming, cooling, packing, monitoring, pressurizing and scaling-up are among the things that can be done to, with or from chromatography columns in this selection — and, the manufacturers will hope, buying.

Level-Alert

From Amika Corp. www.amika.com

Keep track of those columns

Level-Alert is a simple unanchored, wireless device with upper- and lower-level indicators that sounds an alarm whenever the liquid drops or rises beyond a certain level. If the user doesn't respond within a certain time, Level-Alert will automatically take action — shut off the pump, for example. Level-Alert can be used in any situation where the liquid level needs to be monitored, in liquid chromatography, for example, or in solvent bottles or waste bottles.

Reader Service No. 100



Alarming news: Amika can keep you on the level.

Chirobiotic V and T — and is available in guard column, narrow bore, analytical, semi-preparative and preparative column sizes.

Reader Service No. 102

750 Fluorometer

From McPherson www.mcphersoninc.com

All-in-one fluorescence unit

Tired of purchasing multiple fluorescence units to handle multiple types of samples such as HPLC effluents and 10 × 10 mm cuvettes? McPherson, specialists in spectroscopy, offer the model 750 fluorometer as a solution. This can handle 10 × 10 mm cuvettes for scanning spectrofluorometry, HPLC flow cells, capillary electrophoresis flow cells and also chemiluminescence flow cells. The design allows light sources and optics to be optimized for UV, Vis and NIR.

Reader Service No. 103

Lissy TLC

From Zinsser Analytic

www.zinsser-analytic.com

Claimed to be faster and cheaper than HPLC

Zinsser's Lissy TLC thin layer chromatography system is automated for use with the large-throughput methods now involved in drug discovery. The time-consuming and critical process of sample application on the TLC-plates is automated. For high throughput of up to 24 TLC plates, the primary samples are positioned in racks on the instrument. Several hundred samples can be processed in a single run. To avoid spreading of the spots, the sample is dried by a controlled flow of nitrogen.

Reader Service No. 104

Pinnacle

From Beckman Coulter

www.beckmancoulter.com

User-friendly chromatography data system

Pinnacle is a new client/server chromatography data system (CDS) designed for ease of use in medium- to large-scale laboratories.

It combines an Oracle relational database with a user interface modelled after Microsoft Outlook. Pinnacle also incorporates Microsoft Visual Basic for Applications as programming language.

Reader Service No. 105

GC-17AP

From Shimadzu

www.shimadzu.com

Entry-level gas chromatography

Based on the established GC-17A, the new GC-17AP gas chromatograph is a lower-cost alternative, but with many of the features of the higher-cost version. Upgrades are available to suit the machine to specific requirements. The GC-17AP can accommodate up to three injectors and four detectors simultaneously.

Reader Service No. 106

CP-3800

From Varian

www.varian.com

Automation, automation, automation

Launching the CP-3800 this March, Varian claimed it was the most automated gas chromatograph available. It incorporates the Chrompack Micro-TCD detector, electronic flow control to minimize carrier gas consumption, and Ethernet to enable a single operator to control an array of instruments.

Reader Service No. 107

HPLC Fraction Collector

From Waters Corporation www.waters.com

Automatic fraction collector for high flow rates

The Waters Fraction Collector II is a stand-alone system designed for HPLC, preparative LC or open column LC applications. It is compatible with flow rates as high as 300 ml per min and accommodates a wide range of adapter racks and related collection



Waters' flow: a collector for 300 ml per min.

MICRO+

From Chromacol www.chromacol.com

Microsampling made manageable

MICRO+ was developed for analysts working with very small samples, down to 4 µl. MICRO+ is a one-piece assembly of a 300-µl glass tapered insert and a 12 × 32 mm 2-ml borosilicate glass vial. The combination is compatible with all the major chromatography autosamplers which use a standard 2-ml, 12 × 32 mm crimp top vial.

Reader Service No. 101

Chirobiotic R

From BAS www.bas.co.jp

A column targeting anionic chiral molecules

This new LC column, Chirobiotic R, is based on the glycopeptide ristocetin A. It is built to operate in all solvent systems — reversed, normal and polar organic mobile phases. It offers enantioselectivity for a broad range of solutes. The ristocetin A is covalently bound to silica, to target the separation of anionic chiral molecules in particular, and also di- and tripeptides, α-hydroxy acids, substituted aliphatic acids, aromatic esters, chiral alcohols and secondary and tertiary amines. Chirobiotic R joins its sister columns —



Say R: BAS's new columns use ristocetin A.

vessels including microtitre plates, micro-centrifuge tubes, scintillation vials, test tubes, bottles and flasks and the ubiquitous 2-ml vials. An adapter rack allows users with the Waters Alliance HPLC System to collect fractions automatically into two sample carousels each holding 24 2-ml sample vials. **Reader Service No. 108**

PL-GPC 220/120

From Polymer Laboratories

www.polymerlabs.com

Raising the temperature

At PittCon this year Polymer Laboratories launched two new, high-temperature GPC/SEC chromatographs, the PL-GPC 220, for ultra-high temperatures of up to 220°C, and the PL-GPC 120, for medium- to high-temperature operation up to 120°C. New features on the PL-GPC 220 include full PC control, intuitive graphical user interface, Windows compatibility and the Hewlett-Packard 1100 Series isocratic pump. **Reader Service No. 109**

Buffer concentrates

From Hamilton

www.hamiltoncomp.com

Buffer preparation — just add water

These mobile-phase concentrates simplify the preparation of ion chromatography buffers. Buffers for PRP-X100 (anions) and PRP-X200 (cations) columns are prepared by adding water to one of two concentrates. The first, on dilution with water, makes 2 litres of 4.0 mM *p*-hydroxybenzoic acid pH 8.5 with 2.5% methanol. This mobile phase is used with the PRP-X100 column to separate the eight common anions (fluoride, carbonate, bromide, chloride, nitrite, nitrate, phosphate, sulphate). The second concentrate makes 2 litres of 4.0 mM nitric acid in 30% methanol, and with a PRPX200 column can separate monovalent cations (lithium,

sodium, potassium, ammonium). Other buffer concentrations are available to order. **Reader Service No. 110**

Column chiller/heater

From Torrey Pines Scientific

www.torreyпинesscientific.com

Temperature made controllable

The EchoTherm HPLC Column Chiller/Heater Model C030, say the makers, has the largest, easiest to access chamber available. Temperature range is 4.0 °C to 70.0 °C readable and settable to 0.1 °C. The chamber is 15.625 inches tall, 6.385 inches wide and 5.125 inches deep. The chamber door opens down the front, and columns, valves and fittings are mounted on the inside of the door for ready access. **Reader Service No. 111**

Biologic Duo-Flow

From Bio-Rad

www.bio-rad.com

Putting the pressure on proteins

Bio-Rad's new chromatography system features high pressure and flow rates for increased flexibility in protein and peptide purifications. The BioLogic Duo-Flow is a binary gradient system, which can be used for both low- and medium-pressure work. It is capable of 10 ml per min at 3,500 p.s.i. and, with two easily interchangeable pumps, can also run at 40 ml per min at 1,000 p.s.i. for higher sample throughput. There are no metallic components in the fluid flowpath. **Reader Service No. 112**

Nitrogen generator

From Whatman International

www.whatman.com

Pure N² simple

This analytical nitrogen generator is tasked to produce purity levels of dry nitrogen of

99% to 99.9%, for supplying LC/MS instruments as well as for solvent evaporation. To operate the Type N2-2010 generator it only requires to be connected by a 0.5-inch inlet connection to a supply of compressed air ranging from 4.2 to 10 barg to produce a maximum flow rate of nitrogen of 62 litres per minute. The unit is 610 mm wide, 510 mm deep and 1,750 mm high. Weight is 114 kg, there are no electrical connections, and maintenance involves only an annual change of prefilters, a carbon tower and particulate final filter. **Reader Service No. 113**

NovaPrep

From Hitachi

www.hitachi.com

Semipreparative HPLC systems

The NovaPrep 200 HPLC system runs at a flow rate range of 1–200 ml per min while the NovaPrep 800 ranges from 8 to 800 ml per min. Both systems are totally inert. Isocratic, step gradient, low-pressure gradient and high-pressure gradient configurations are available. A wide range of columns and sorbents are available. Up to three data collection channels are supported, and Hitachi offers UV, UV/Vis, pH, conductivity, temperature, RI and other detectors. NovaPrep's integrated design is intended to make scale-up simple. **Reader Service No. 114**

Sensor chips

From TRACE Biotech AG

www.trace-ag.de

Electrochemical detection for HPLC

TED 4020 detector systems from TRACE make use of a new type of sensor chip, made by layering thin coatings on a smooth substrate to be economical with the expensive detector material. The resulting electrodes are simply changed when they wear out. The time it takes for the electrodes to build up to full intensity is also claimed to be much shorter than for conventional electro-

ADVERTISEMENTS

Homogenizers?

- Tissue-Tearor™
- BeadBeater™
- Mini-BeadBeater™
- BioHomogenizer™

Check Us Out

www.biospec.com

BioSpec Products, Inc.
Ph/Fax 918-336-3363
e-mail info@biospec.com

New Probe Sizes
0.45cm!

READER ENQUIRY NO. 54

Detect residual DNA with Wako's

DNA Extractor Kit

or Isolate Genomic DNA with

DNA Extractor WB Kit

Both use the NaI Method
No Phenol No Chloroform

Wako

Wako BioProducts
Wako Chemicals USA, Inc.
1600 Bellwood Road
Richmond, VA 23237
Telephone: (804) 271-7677
Facsimile: (804) 271-7791
(800) 992-9256
bioproducts@wakousa.com

READER ENQUIRY NO. 55



HPLC with chips: TRACE's new sensor.

chemical cells. TRACE offers this module in various materials (including PEEK, polycarbonate) and various geometric cell designs.
Reader Service No. 115

HP 1100 Series

From Hewlett Packard www.hp.com
Multi-wavelength detection in HPLC

A high-sensitivity multiwavelength detector has been added to the HP 1100 series of HPLC systems to detect a wide range of wavelengths, from low UV to the visible range. The diode-array-based optical design provides high sensitivity for trace-level quantification. Simultaneous detection of multiple signals improves selectivity as well as sensitivity. The detector fits neatly in an HP 1100 Series tower.

Reader Service No. 116

FlashMaster

From Jones Chromatography
www.jones-chrom.co.uk

Evolution in flash chromatography

FlashMaster is a complete, software-controlled system with automated multiple-sample operation and intelligent fraction collection with sample tracking. It features on-line UV detection that allow separations to be seen, and the use of a pumping system is claimed to ensure reproducible flow rates. FlashMaster is a three-stage evolution from traditional flash chromatography with on-line UV detection as the first stage, linear gradient elution as the second, and targeted fraction collection the third.

Reader Service No. 117

DH-2000-HPLC

From Ocean Optics www.OceanOptics.com

Monitor multiple wavelengths simultaneously

DH-2000-HPLC is an integrated flow-cell light source that functions as a compact, high-sensitivity detector for HPLC analysis.



Ocean waves: the multi-λ DH-2000.



HP terms: Hewlett Packard's detector module.

The DH-2000-HPLC offers an optical detection method with the ability to capture complete spectral information. It can simultaneously monitor multiple wavelengths, as well as capture the spectra of analytes as they pass through the detector.

Reader Service No. 118

HyPURITY ADVANCE

From ThermoQuest Hypersil
www.thermoquest.com

Just going through a phase

This new HPLC phase, named HyPURITY Advance, is aimed especially at the drug discovery and combinatorial chemistry markets. Based on the HyPURITY silica base, the polar embedded group features a C8 grouping and amide linkage chemistry. This combination allows rapid separation of a wide range of pharmaceutical compounds, especially those exhibiting strongly acidic and/or basic functionality. The phase is available as both traditional 5-μm and high-speed 3-μm particles. Columns are available in varying lengths and from 320 μm to 21.2 mm internal diameter.

Reader Service No. 119

DuraGuard

From J&W Scientific www.jandw.com

GC column that traps sample residues

DuraGuard columns, latest additions to the J&W Scientific range for high-resolution capillary GC, are equipped with a built-in guard column to trap nonvolatile sample residues. Retention gaps are used to focus the injected sample so as to improve the peak shapes when using on-column and splitless injectors. The guard column, retention gap or transfer line and the analytical column are made with a single, continuous piece of fused silica tubing, thus eliminating the need for a union to attach the deactivated fused silica tubing to the analytical column.

Reader Service No. 120

DAD 540

From Bio-Tek www.biotech-kontron.com

Fibreoptic detection

Bio-Tek's DAD 540+ and DAD 545 HPLC

diode array detectors function in the UV and UV-Vis ranges. A range of analytical flow cells is available to cover all HPLC applications, including analytical and microbore. Long fibre optics allow the flow cell to be positioned outside the detector module without the risk of deterioration of chromatographic and spectral performance.

Reader Service No. 121

FAST GC columns

From SGE Europe Ltd www.sge.com

FAST by name, and extremely rapid by nature

Milton Keynes-based company SGE has introduced a new range of columns, called FAST GC. The makers claim ultra-high efficiencies, producing extremely quick analysis with baseline resolution. The columns are short (10 m) with an inner diameter of 0.1 mm, and are coated with thin films of SGE low-bleed phases. They are especially suited to labs screening for pesticides, drugs and polychlorinated biphenyls.

Reader Service No. 122

GSTrap

From Amersham Pharmacia Biotech
www.apbiotech.com

An affinity for purification

GSTrap affinity columns are designed for the one-step purification of glutathione S-transferase fusion proteins and other glutathione-dependent proteins. The fusion proteins are eluted under mild, non-denaturing conditions which preserve protein antigenicity and functionality, and can be purified directly from bacterial lysates. The column matrix, Glutathione Sepharose 4 Fast Flow, is also available as pre-swollen loose media in a 95-ml package.

Reader Service No. 123

These notes are compiled in the Nature office from information provided by the manufacturers. For more details, fill in the reader service card bound inside the journal.



The matrix: or you can buy GSTrap as a column.

Green chemistry puts down roots

Industry is discovering that 'green' approaches to chemical processes are not only beneficial to the environment but can boost profits too. It's fertile ground for collaboration between academic and industrial scientists.

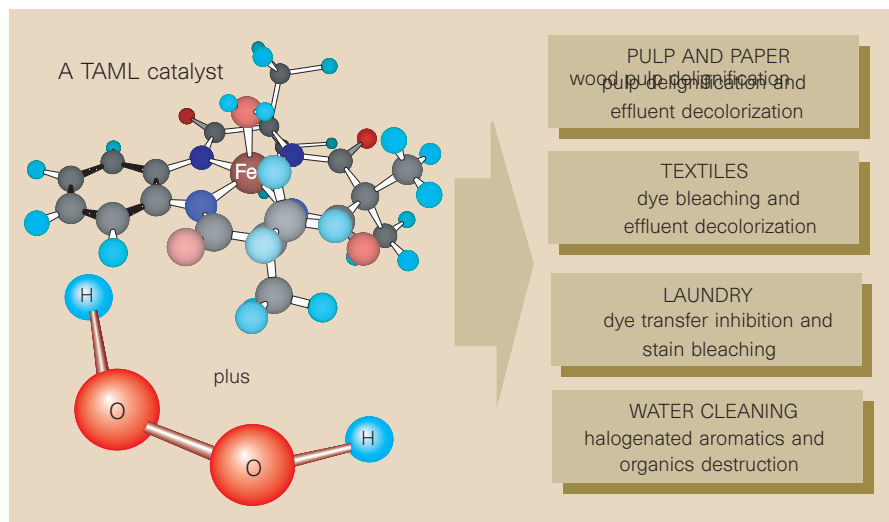
Brendan Horton

The journal *Chemical and Engineering News* downsized its list of top chemical companies from 100 to 75 last spring: mergers and acquisitions had taken their toll of the annual rankings. Although the economic pressures on this mature industry seem clear, 'green' chemistry approaches, which have previously been spurned, are now spurring innovation and improving industrial competitiveness. The applied nature of green chemistry has resulted in a great deal of collaboration between government, industry and academic institutions, as well as between chemists, engineers, biologists and business people. Such collaborations have resulted in the formation of green chemistry centres and institutes, and may bring new norms for how to do research and development.

Green chemistry (or 'sustainable' chemistry, which is the preferred adjective in Europe) is showing that the ability to rethink traditional chemistry processes and to design closed-system manufacturing processes is not only beneficial to the environment but also makes economic sense for the industry. Green chemistry is creating tools that allow industry to move towards the goals of 'industrial ecology'. Paul Anastas, chief of the industrial chemistry branch at the US Environmental Protection Agency (EPA), and known to some as the father of the term green chemistry, describes it as the design of chemical products and processes that reduce or eliminate the use and generation of hazardous substances.

According to Terry Collins, winner of this year's academic prize in the US President's Green Chemistry Challenge, "chemistry is the key science of sustainability". Collins, an inorganic chemist at Carnegie Mellon University in Pennsylvania, emphasizes three areas where significant innovation is needed: solar-to-chemical energy conversion, toxicity reduction, and biomass conversion as a source of long-term feedstocks for the chemical industry.

Industry has an important role in this movement. Collins draws attention to the fact that there are four industrial awards in the Green Chemistry Challenge and only one academic award at the annual Green Chemistry and Engineering conference, which was held last month in Washington, DC. "The EPA and the organizations behind the awards are recognizing developments within the chemical industry where people are making major improvements to existing products or producing something that



Science of sustainability: Collins and his team from Carnegie Mellon University won an award in the US Green Chemistry Challenge for developing tetraamido-macrocyclic ligand (TAML) catalysts. They could have an impact in a range of industrial processes.

improves biomass transfer or reduces toxicity," he says. But Collins is troubled that there is relatively little work being done on solar-to-chemical energy transformation: "We should have an enormous amount of work going on in that area, and we don't."

In contrast to earlier environmental movements, in which protection was thought to be a cost burden to industry, many areas of green chemistry are now being driven by industry, says Anastas. As far as he is aware, this is the first time that, through basic science, it is being shown that environmental protection can be profitable to industry.

When the EPA initiated its push into green chemistry in the early 1990s, it designed programmes to attract industrialists, academics and environmentalists. This was done by including as founding members the Chemical Manufacturers Association, chemical companies, the American Chemical Society, the National Academy of Sciences, the Council for Chemical Research and the Environmental Defense Fund.

The EPA is mainly a regulatory agency, but in the case of the green chemistry initiative it is acting more like a grants agency and a facilitator of technological development. Its programmes in green chemistry "do not have a regulatory bone in their body," says Anastas. "Green chemistry approaches are being used to find ways of meeting economic and environmental goals without having to regulate."

The academic community's participation is based on the fundamental scientific challenges to be overcome in such areas as

solvents, benign processing, toxicity, catalysis, bio-based synthesis and processing, and solar-to-chemical conversion. Through this green movement, a new culture of collaboration seems to be developing to do research and engineering, raise money, and make technology development more efficient.

Successful collaboration

For Ken Seddon, a professor of chemistry at the Queen's University of Belfast in Northern Ireland, linking the fundamental chemistry to industry is very important: "We are very much concerned with having an applied direction to our research." His research area, ionic liquids, provides many opportunities for doing fundamental work as well. "Everything that we've learned or guessed about chemistry is based on observations in a molecular medium," says Seddon. But this new class of solvents offers very different kinetics and thermodynamics, as well as the potential for improved product yields and selectivity.

For example, he says, the Friedel-Crafts reaction conventionally takes 8 hours at 80 °C and produces 80% yield with a mixture of isomers. Using ionic liquids, the same reaction takes 30 seconds at 0 °C and has a 98% yield of a single isomer. That is impressive in itself, but there is also a potential significant environmental benefit with ionic liquids. They should prove to be an effective substitute for an environmentally troublesome class of solvents common to many synthetic pathways, namely volatile organic compounds or VOCs.

Ionic liquids remain in a liquid state

through the temperature range -97°C to 200°C , they have no effective vapour pressure, and they are relatively cheap and easy to prepare. Exploratory work in Seddon's laboratory (in collaboration with BP Chemicals and Unilever Port Sunlight Research Laboratory) has demonstrated that a wide range of catalysed organic reactions occur in room-temperature ionic liquids, including oligomerizations, polymerizations, alkylations and acylations. These are serious candidates for commercial processes.

According to Seddon, about half the work he was doing with his industrial collaborators was confidential; the rest was generic research on ionic-liquid solvent systems. A meeting between nine of his industrial collaborators soon revealed that they would be happy to share the generic half of their research on ionic liquids with each other. This resulted in the formation of the Queen's University ion liquids laboratory or QUILL, with member companies putting up £20,000 (US\$32,000) each year for membership.

Not all the work for QUILL goes on in the university's laboratories. "The companies' workers will work in our labs; our workers will work in their labs. It's a genuine collabora-

tion—a two-way street," says Seddon. Not only do these collaborations bring together academics and industrialists, they also draw on a variety of scientific talent, including chemical engineers, and physical, inorganic, organic, pharmaceutical, and computational chemists. The 17 companies involved reap better returns by sharing the generic information, as they can then focus more on their specialized interests.

Multidisciplinary approaches

Not far removed from Seddon, in either his green philosophy or his desire to bring industrialists and academics together, is Chris Adams, an inorganic chemist with 24 years' industrial experience at Unilever. A couple of years ago, when Unilever sold its chemical interests, Adams decided to create a new kind of research organization, which became the Institute of Applied Catalysis (IAC).

The institute spun out of the UK government's Foresight programme, which, according to Adams, "was an attempt to develop consensus and enthusiasm for what science and technology can do for quality of life and wealth creation". The programme pulled together several thousand senior

technical people from UK industries and universities and organized them into panels by industry sector. At the time, says Adams, "we had a lot of good fundamental science in catalysis, but we didn't have vehicles to do the first stages of engineering". As a top priority, the Foresight chemicals panel recommended the formation of a national institute of applied catalysis to close this gap.

The original idea was to build new premises. But the panel realized that this might cost £20 million, says Adams, so instead a virtual consortium was set up, with Adams as director. "I wouldn't say it's completely run by myself from my lap-top computer, but it's close to that." There are 15 member companies, ranging from such heavy hitters as BP/Amoco, ICI and Shell, through Johnson Matthey, Air Products, British Nuclear Fuels, British Gas Technology and Astro Zeneca, right down to some quite small ones. The agreement, according to Adams, was to build a community of scientists and technologists from industry and universities with people trained in chemistry, chemical engineering, materials, modelling and computing.

"We insisted that research projects be multidisciplinary and there has been tremendous response to this," says Adams. Well over £2 million has been put into the universities, and the academics have responded favourably. "They get huge industrial input into the projects, with lots of steering and management. We put in extra training for the students on things like intellectual property, project management and team working." Adams feels that IAC is one of the few groups that are getting multidisciplinary approaches and teamwork to function effectively.

It is impossible to do industrial chemistry without bringing together a variety of disciplines, says Adams. With that in mind, IAC has set up a formal education programme, the first leg of which is aimed at postgraduates. The courses comprise industrial case studies, and are designed to be highly participatory. "We get 30 young chemists and engineers together for a week, give them some real problems to work on, and get them working through the night in multidisciplinary teams," says Adams. They work over several different dimensions, he says, from taking lab results and extrapolating them to the engineering side, to factoring in economic and environmental effects.

Cleaner and cheaper

When asked what employers look for when hiring people to do green chemistry, James Bashkin, associate editor of the newly launched *Journal of Green Chemistry*, says: "I would look for the best trained person in organic or inorganic chemistry with the willingness to be non-traditional and the ability to interact with people from other disciplines."

Further information on the web

Center for Green Manufacturing

<http://bama.ua.edu/~cgm/>

Ken Seddon

<http://www.ch.qub.ac.uk/staff/personal/krs/krs.html>

QUILL

<http://quill.qub.ac.uk/>

Ionic Liquids Review

<http://www.ch.qub.ac.uk/resources/ionic/review/review.html>

Robin D. Rogers

<http://bama.ua.edu/~rdrogers/>

Terry Collins

<http://www.chem.cmu.edu:80/Collins.html>

Chris Adams

<http://www.iac.org.uk/>

James Bashkin

<http://wunmr.wustl.edu/Faculty/Bashkin/index.html>

EPA Presidential Green Chemistry Challenge

<http://www.epa.gov/greenchemistry/PresidentialGreen>

EPA green chemistry page

<http://www.epa.gov/greenchemistry/index.htm>

Journals and reports

The Journal of Green Chemistry

<http://www.rsc.org/is/journals/current/green/greenpub.htm>

Clean Products and Processes

<http://link.springer.de/link/service/journals/10098/index.htm>

RAND report on environmental technology

<http://www.rand.org/publications/MR/MR1068/>

International Congress of Chemistry and

Environment

<http://www.chemenvron.com/icce1.html>

Networks, centres and institutes

Green Chemistry Network

<http://www.chemsoc.org/gcn/index.htm>

Green Chemistry Institute

<http://www.lanl.gov/greenchemistry/>

Italian Group of Catalysis

<http://www.fci.unibo.it/gic/>

Netherlands Institute of Catalysis Research

<http://www.nl-knowhow.org/organisations/NIOK.html>

Dechema

http://www.dechema.de/englisch/fue/nice/pages/f_index.htm

Catalyse et chimie des matériaux divisés

<http://www.agro.ucl.ac.be/cata/nice.html>

Fifth Framework programme: energy,

environment and sustainable development

<http://www.cordis.lu/eesd/home.html>

Chemistry societies

Royal Society of Chemistry

<http://www.rsc.org/>

American Chemical Society

<http://www.acs.org/>

ACS careers services

<http://www.acs.org/careers/welcome.htm>

Chemical Manufacturers Association: 1999

Responsible Care Conferences, ChemRAWN

<http://iupac.chemsoc.org/symposia/conferences/chemrawn/chemrawnIX.html#2>

Gordon Research Conference

<http://www.grc.uri.edu/proURLgrams/1999/green.htm>

Chemistry conferences

<http://www.rsc.org/lap/confs/confshome.htm>

Before joining the faculty at Washington University, Saint Louis, Bashkin developed a reaction for Monsanto's rubber chemicals division that eliminated a step in a complex process. This process produces 200 million tonnes a year of an antioxidant rubber additive, which keeps rubber from ageing.

According to Bashkin, the use of this reaction has changed the global economy in the business. "Not only did we eliminate a pollution problem, but we made the manufacture more economical." For this work, Bashkin and his colleagues from Flexsys (a joint venture of Monsanto and Akzo Nobel) received one of last year's presidential Green Chemistry Challenge awards.

Traditionally, academic chemistry tends to be an isolating experience, says Bashkin. "You have to do your own work and you're dependent on yourself and your adviser. It's not necessarily good for one's career to be involved in collaborations early on." But in industry, he says, nothing is accomplished without collaboration.

As a PhD student, you need to make sure that there is a definable body of work that is your own, which tends to work against collaboration. It is hard for untenured faculty members to collaborate, says Bashkin, owing to the insistence that you develop your own professional identity. "I don't think that's necessarily good for science," he says. "In the modern world, we can accomplish a lot more if we allow ourselves to be influenced by others."

The Center for Green Manufacturing at

the University of Alabama has won funding from the US Department of Energy's Office of Industry. It was formed to build industry leadership in the use of alternative reaction and separation media in Alabama. According to Robin Rogers, a professor of chemistry and director of the centre, workshops will introduce plant managers to new technologies, such as room-temperature ionic liquids as new media for carrying out reactions.

The future is green

Rogers believes that, unless greener technologies can be developed for chemical manufacturers and pharmaceutical companies, there is no long-term future for those industries in the United States. Companies will simply move their plants to countries with fewer regulations and lower costs for waste disposal.

"When I was in graduate school in the organic chemistry lab, I was graded on how much product I had. That's how I was trained to think: maximize your yield." But, says Rogers, there are other aspects that must be taken into consideration now and be taught to students. "We can see that the future lies in team approaches to problem solving and in thinking about processes from a different perspective." He asks: "How many people in science do you know who work with business people when they develop the idea that they want to pursue?" To address this, Rogers has brought together colleagues from the school of business, the college of engineering, and



Team player: Rogers at the site of new research facilities for the Center for Green Manufacturing.

several of the sciences to expose students to the complete cycle of process development.

According to Rogers, there are many institutional and cultural barriers to overcome, and students must understand that there are other factors besides chemistry that go into solving a problem, whether in economics or engineering. "But how can my students learn the true engineering needs by being taught by just me?" Rogers believes that, if students are involved in projects that expose them to these other perspectives and requirements, they will be better prepared for the real world than chemists who are not exposed to them until they reach industry. □

*Brendan Horton is a freelance journalist.
e-mail: brenhorton@aol.com*

Data explosion fuels search for drugs

Potter Wickware

In the data-driven world of drug discovery, chemists should be comfortable not only with new perspectives on their data but also with new ways of relating to their fellow scientists. So says Steve Kaldor, director of research at Eli Lilly, a large pharmaceutical company which relies increasingly on combinatorial chemistry for its efforts at drug discovery. "Because many biological targets have relatively poor validation states as true drug candidates, it's important to be able to sift through assay data from many simultaneous projects," he says.

And, because chemists are increasingly likely to be members of multidisciplinary teams, perhaps as leader in one project and in a supporting role in another, teamwork skills are also more important than ever. "Chemistry used to be much more linear, both as it was taught in school and practised in industry, but this is rapidly changing. The lone wolf is a rarity now as a successful model for a drug-discovery chemist," says Kaldor.

Research strategies are also evolving rapidly. Dave Hangauer, of the Department



Kaldor: 'develop data-crunching skills'.

of Medicinal Chemistry at the University of Buffalo, New York state, designs and tests protein kinase inhibitors using small-molecule libraries produced both by solution and solid support synthesis. He explains that, because the number of possible compounds in molecular space is far too large to screen, workable combinatorial libraries depend on astute initial hypotheses to steer the output towards a manageable subset.

"Structure-based design of combinatorial libraries is a nice example of hypothesis-driven science combined with data-driven science," he says. For example, when a library of ligands modelled to contain tight binders is assayed experimentally in a high-throughput screen, the molecular modelling assumptions can be quickly tested and hypotheses suggested. "The underlying science of predicting the free energy of binding

in biological buffers is poorly developed, but combinatorial methods allow experiments to be pointed in promising new directions."

The new techniques also highlight fruitful areas of investigation at the interface of areas that may have been viewed previously as distantly related, such as organic chemistry and genomics. Carolyn Bertozzi, a biological chemist at the University of California, Berkeley, uses combinatorial library screening to explore enzymes involved in carbohydrate biosynthesis. She says: "In drug discovery, small-molecule modulators and genomic data are intertwined, so any chemist who desires a career in research at the chemistry/biology interface would benefit from training in genomic analysis."

Because many drug targets are enzymes from superfamilies, it is a major challenge to develop small-molecule inhibitors that are selective for one individual within a large related group of targets. "If sequence homologies could be correlated with potent and specific inhibitor structures, useful information could be obtained to guide drug discovery and elucidate biological pathways."

The library and high-throughput

approach, once confined to organic molecules for drug discovery, has quickly outrun its original range. Henry Weinberg, research director at Symyx, a materials science company in Santa Clara, California, uses combinatorial methods to explore the entire periodic table for solid-state and solution catalysts, electronic materials such as phosphors, and polymers. He says: "The spirit of rapid synthesis, followed by rapid parallel or serial screening, followed by data management, is the same as in the biological realm, but with us the particulars are quite different." The screening of inorganic materials can be particularly challenging because the performance characteristics sought may be unlike those in a typical screen for receptor binding. Accomplishing this esoteric work requires the efforts of many disciplines, from physical chemists to condensed-matter physicists to mechanical and software engineers.

As the nature of the data and the problems have changed, so has the manner in



Weinberg: 'screening offers challenges'.

which research is organized. Bertozzi says "collaborative projects are more popular than ever, and I see a trend towards more open sharing of resources to solve bigger problems. Students who participate in collaborative projects will have superior job prospects."

Hangauer says that the traditional model of a hierarchically organized lab with a research project directed by a single principal investigator (PI) is waning. "In the ascendant multidisciplinary environment, students will be obliged to have broader interests, abilities and training while still maintaining a strength in some particular area, such as organic synthesis," he says. The requirement will extend to senior members of the team. And for decisions on who should be given tenured posts, a revision in the way productivity is evaluated when research originates from a multiple PI team is probably inevitable. Bertozzi agrees: "Tenure decisions still focus too much on the delineation of credit, which detracts from the power of the collaboration."

Technology plays a curiously ambivalent role in the modern chemist's toolkit. On the one hand, laboratory robotics, nanoscale-array fabrication for large-scale synthesis, computer analysis, and database design and management issues are almost certain to penetrate the working chemist's laboratory over the next decade. "And yet a lot of the technology simply won't stick," says Kaldor. "Combinatorial chemistry itself was not a major force in the pharmaceutical industry even eight years ago and, while it is likely to be around for a long time to come, what it

looks like eight years from now may well be unrecognizable by today's standards."

There's a real threat, he adds, of scientists being seduced by new technology and diverted away from the goals of their research. "In the same way you'd classify drugs as leading to targets or not, you can classify technology as being useful or useless. The scientist who can pragmatically evaluate a new tool from a chemistry perspective I put at a high value for recruiting and retaining as an employee," says Kaldor. The basic skills have not changed in any radical way with the advent of combinatorial libraries or other technologies. Central to any career will remain the ability to design an experiment that will work, and to look at data in a new way to recognize some important feature that may have been hidden.

Skill shortage

Amir H. Hoveyda, at the Chemistry Department at Boston College, Massachusetts, uses high-throughput screening to investigate metal-ligand combinations with the goal of introducing new reagents and unusual modes of reactivity to organic chemistry. He agrees that particular skills are less important than more general ones. "Particular skills change and vary from place to place, and good people can always pick up what may be required." He says that a basic skill in short supply is organic synthesis.

"In the 1980s many synthetic organic chemists defected to biology, so today there are few people in their forties and fifties who are training them. Meanwhile, the employment demand is growing, caused in part by the fact that combinatorial chemistry has provided pharmaceutical companies with many leads." In his view, the best basic graduate training is in synthetic organic chemistry, because "if you know how to make molecules and how they behave, that knowledge can be applied to almost any branch of chemistry."

Future areas of growth are likely to

include solid supported synthesis and other technologies that will allow longer synthetic sequences than solution chemistry, according to Hangauer. There is also room for improvement in the analysis and measurement of molecular diversity. 'Chemical evolution' to expand the diversity of libraries represents another area ripe with promise. This relies on approaches such as those of Hangauer's colleague Alexey V. Eliseev, who is developing an evolutionary optimization of small-molecule synthetic libraries driven by the biological target.

Bertozzi says that, although synthetic chemistry is of primary significance, database management and data-mining algorithms also need development so that underlying principles latent in large volumes of data can be extracted with more reliability.

Bertozzi: 'get trained in genomic analysis'.

None of the scientists interviewed believes that an end to the growth in chemistry is anywhere in sight, despite a *fin-de-siècle* pessimism about the prospects for science that one sometimes hears from physicists. Hangauer comments that, on the contrary, "chemistry is poised to move rapidly forward with much greater efficiency and to uncover new principles in areas such as non-covalent interactions".

The job market is strong, says Kaldor. At his company, "a recruiting effort similar in magnitude to the largest we've ever made" is under way. He thinks that a big pharmaceutical company is a great place for a chemist to start a career. "The newcomer can be mentored by people who know the business. The learning one gets is transferable. There's more opportunity to move across the various stages of the drug-discovery process. If they decide later that discovery is not for them, they can move into a development career without uprooting their career and family."

Students ought to pick a core focus area that has broad use, such as organic synthesis, and add to that specific areas of interest such as medicinal chemistry or chemoinformatics. It is important to combine strong, universally applicable fundamentals with some breadth in the training. Choose an institution with a good reputation and a broad selection of courses and research opportunities. Choose a PI with a strong interest in the welfare of his or her students and who provides exciting research opportunities.

Finally, advises Bertozzi, "start your career in research as soon as possible. Get a job in a lab to see how the process works. Get trained in overlapping disciplines to position yourself for a career at the cutting edge". □

Potter Wickware is a science writer in California.
e-mail: wickware@worldnet.att.net

Web reading

Symyx Technologies: combinatorial chemistry for chemical and electronic industries

<http://www.symyx.com/>

Carolyn R. Bertozzi: bio-organic chemistry

http://www.cchem.berkeley.edu:80/~chemgrad/grad_program/faculty/bertozzi/bertozzi.html

Dave Hangauer

http://wings.buffalo.edu/academic/departments/pharmacy/mch/public_html/fac/hangau.html

Combinatorial chemistry

<http://www.netsci.org/Science/Combichem/feature02.html>

Amir H. Hoveyda, Boston College

<http://chemserv.bc.edu/Department/Faculty/hoveyda/>

Eli Lilly & Co.

<http://www.lilly.com/>